10/537798
Rec'd PCT/PTO 07 JUN 2005
16. Dez. 2003
PCT/EP03/14344
#2

Europäisches Patentamt | European Patent Office | Office européen des brevets

REC'T 02 MAR 2004
WIPO PCT

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. | Patent application No. | Demande de brevet n°

02028275.2

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

BEST AVAILABLE COPY

Anmeldung Nr:
Application no.: 02028275.2
Demande no:

Anmeldetag:
Date of filing: 16.12.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Strasse 28
37079 Göttingen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Eip75B, Cdk4, or CG7134 homologous proteins involved in the regulation of energy homeostasis

In Anspruch genommene Prioriät(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

A61K38/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SI SK

16. Dez. 2002

WEICKMANN & WEICKMANN

Patentanwälte

European Patent Attorneys · European Trademark Attorneys

EPO - Munich
61
16. Dez. 2002

DIPL.-ING. H. WEICKMANN (bis 31.1.01)
DIPL.-ING. F. A. WEICKMANN
DIPL.-CHEM. B. HUBER
DR.-ING. H. LISKA
DIPL.-PHYS. DR. J. PRECHTEL
DIPL.-CHEM. DR. B. BÖHM
DIPL.-CHEM. DR. W. WEISS
DIPL.-PHYS. DR. J. TIESMEYER
DIPL.-PHYS. DR. M. HERZOG
DIPL.-PHYS. B. RUTTENSPERGER
DIPL.-PHYS. DR.-ING. V. JORDAN
DIPL.-CHEM. DR. M. DEY
DIPL.-FORSTW. DR. J. LACHNIT

Unser Zeichen:
29583P EP/WWld

Anmelder:
DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Straße 28

37079 Göttingen
DEUTSCHLAND

Eip75B, Cdk4, or CG7134 homologous proteins involved in the regulation of energy homeostasis

---

EPO - Munich
61

16. Dez. 2002

Eip75B, Cdk4, or CG7134 homologous proteins involved in the regulation of energy homeostasis

Description

This invention relates to the use of nucleic acid sequences encoding Eip75B, Cdk4, or CG7134 homologous proteins, and the polypeptides encoded thereby and to the use thereof or effectors of Eip75B, Cdk4 or CG7134 homologous proteins in the diagnosis, study, prevention, and treatment of metabolic diseases or dysfunctions, for example, but not limited to, such as metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones.

There are several metabolic diseases of human and animal metabolism, eg., obesity and severe weight loss, that relate to energy imbalance where caloric intake versus energy expenditure is imbalanced. Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight, frequently resulting in a significant impairment of health. Obesity may be measured by body mass index, an indicator of adiposity or fatness. Further parameters for defining obesity are waist circumferences, skinfold thickness and bioimpedance (see, inter alia, Kopelman (1999), loc. cit.). It is associated with an increased risk for cardiovascular disease, hypertension, diabetes mellitus Type II, hyperlipidaemia and an increased mortality rate. Besides severe risks of illness, individuals suffering from obesity are often isolated socially.

Obesity is influenced by genetic, metabolic, biochemical, psychological, and behavioral factors and can be caused by different reasons such as non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure. As such, it is a complex disorder that must be addressed on several fronts to achieve lasting positive clinical outcome. Since obesity is not to be considered as a single disorder but as a heterogeneous group of conditions with (potential) multiple causes, it is also characterized by elevated fasting plasma insulin and an exaggerated insulin response to oral glucose intake (Koltermann J., (1980) Clin. Invest 65, 1272-1284). A clear involvement of obesity in type 2 diabetes mellitus can be confirmed (Kopelman P.G., (2000) Nature 404, 635-643).

Hyperlipidemia and elevation of free fatty acids correlate clearly with the 'Metabolic Syndrome', which is defined as the linkage between several diseases, including obesity an insulin resistance. This often occurs in the same patients and is a major risk factor for development of Type 2 diabetes and cardiovascular disease. It was suggested that the control of lipid levels and glucose levels is required to treat Type 2 Diabetes, heart disease, and other occurances of Metabolic Syndrome (see, for example, Santomauro A. T. et al., (1999) Diabetes, 48(9):1836-1841).

The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin, VCPI, VCPL or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations are not known. In addition, several single-gene mutations resulting in obesity have been described in mice, implicating genetic factors in the etiology of obesity (Friedman J. M. and Leibel R. L., (1992) Cell 69(2): 217-220). In the ob mouse a single gene mutation (obese) results in

profound obesity, which is accompanied by diabetes (Friedman J. M. et. al., (1991) Genomics 11: 1054-1062).

Therefore, the technical problem underlying the present invention was to provide for means and methods for modulating (pathological) metabolic conditions influencing body-weight regulation and/or energy homeostatic circuits. The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity. The present invention discloses specific genes involved in the regulation of body-weight, energy homeostasis, metabolism, and obesity, as well as related diseases such as diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones. In particular, the present invention describes the human Eip75B, Cdk4, or CG7134 homologous genes as being involved in those conditions mentioned above.

Nuclear receptor subfamily 1, group D, member 1 (NR1D1) and member 2 (NR1D2) are members of the nuclear receptor superfamily of ligand-activated transcription factors. NR1D1 is a thyroid/steroid hormone receptor and functions as a transcriptional regulator. The NR1D1 gene encodes the orphan receptor Rev-ErbA alpha, and NR1D2 encodes Rev-ErbA beta (Koh Y. S. and Moore D. D., (1999) Genomics 57(2):289-292). Rev-ErbA alpha (Rev-Erb) mRNA levels increased during the differentiation of 3T3-L1 cells into adipocytes. Rev-Erb was similarly induced in the related 3T3-F442A cell line but not in nondifferentiating 3T3-C2 cells. Treatment of preadipocytes with retinoic acid inhibited adipocyte differentiation and also prevented Rev-Erb induction (Chawla A. and Lazar M. A., (1993) J Biol Chem 268(22):16265-16269). Circadian rhythms are generated by a transcription/translation feedback loop

consisting of two limbs, one positive and one negative. The nuclear orphan receptor, Rev-Erb alpha, was identified as a molecular link coupling these two limbs (Alvarez .J. D and Sehgal A., (2002) Dev Cell 3(2):150-152).Rev-ErbA beta shows a high expression in particular in the cerebellum, the dentate gyrus of the hippocampus and pituitary gland of adult rats (Enmark E. et al., (1994) Biochem Biophys Res Commun 1994 Oct 14;204(1):49-56).

Cyclin-dependent kinase 6 (CDK6) and cyclin-dependent kinase 4 (CDK4) are members of the cyclin-dependent protein kinase (CDK) family. CDK family members are highly similar to the gene products of Saccharomyces cerevisiae cdc28, and Schizosaccharomyces pombe cdc2, and are known to be important regulators of cell cycle progression. CDK6 and CDK4 are catalytic subunits of the protein kinase complex that is important for cell cycle G1 phase progression and G1/S transition, these kinases have been shown to phosphorylate, and thus regulate the activity of, tumor suppressor protein retinoblastoma protein (Rb).
CDK4(-/-) mice survived embryogenesis and showed growth retardation and reproductive dysfunction associated with hypoplastic seminiferous tubules in the testis and perturbed corpus luteum formation in the ovary. A majority of CDK4(-/-) mice developed diabetes mellitus by 6 weeks, associated with degeneration of pancreatic islets. (Tsutsui T. et al., (1999) Mol Cell Biol 19(10):7011-7019).

The protein encoded by CDC14 is a member of the dual specificity protein tyrosine phosphatase family. This protein is highly similar to Saccharomyces cerevisiae Cdc14, a protein tyrosine phosphatase involved in the exit of cell mitosis and initiation of DNA replication, which suggests the role in cell cycle control. This protein has been shown to interact with and dephosphorylates tumor suppressor protein p53, and is thought to

regulate the function of p53. Alternative splice of this gene results in 3 transcript variants encoding distinct isoforms.

CDC14A gene is expressed as 1.8- and 4.4-kb mRNAs in all tissues, with the strongest expression in kidney, heart, and skeletal muscle, CDC14B shows ubiquitious expression. In budding yeast, the Cdc14p phosphatase activates mitotic exit by dephosphorylation of specific cyclin-dependent kinase (Cdk) substrates and seems to be regulated by sequestration in the nucleolus until its release in mitosis. Herein, we have analyzed the two human homologs of Cdc14p, hCdc14A and hCdc14B. We demonstrate that the human Cdc14A phosphatase is selective for Cdk substrates in vitro and that although the protein abundance and intrinsic phosphatase activity of hCdc14A and B vary modestly during the cell cycle, their localization is cell cycle regulated. hCdc14A dynamically localizes to interphase but not mitotic centrosomes, and hCdc14B localizes to the interphase nucleolus. These distinct patterns of localization suggest that each isoform of human Cdc14 likely regulates separate cell cycle events. In addition, hCdc14A overexpression induces the loss of the pericentriolar markers pericentrin and gamma-tubulin from centrosomes. Overproduction of hCdc14A also causes mitotic spindle and chromosome segregation defects, defective karyokinesis, and a failure to complete cytokinesis. Thus, the hCdc14A phosphatase appears to play a role in the regulation of the centrosome cycle, mitosis, and cytokinesis, thereby influencing chromosome partitioning and genomic stability in human cells (Kaiser B. K. et al. (2002) Mol Biol Cell 13(7):2289-2300).

Cdc14A phosphatase interacts with interphase centrosomes, and that this interaction is independent of microtubules and Cdc14A phosphatase activity, but requires active nuclear export. Disrupting the nuclear export signal (NES) led to Cdc14A being localized in nucleoli, which in unperturbed cells selectively contain Cdc14B (ref. 1). Conditional overproduction of Cdc14A, but not its phosphatase-dead or NES-deficient mutants, or Cdc14B, resulted in premature centrosome splitting and formation of supernumerary mitotic spindles. In contrast, downregulation

of endogenous Cdc14A by short inhibitory RNA duplexes (siRNA) induced mitotic defects including impaired centrosome separation and failure to undergo productive cytokinesis. Consequently, both overexpression and downregulation of Cdc14A caused aberrant chromosome partitioning into daughter cells. These results indicate thatCdc14A is a physiological regulator of the centrosome duplication cycle, which, when disrupted, can lead to genomic instability in mammalian cells (Mailand N. et al., (2002) Nat Cell Biol 4(4):317-322). CDC14 is involved in the control of the cell cycle (review Oliferenko S. and Balasubramanian M. K. (2001) Curr Biol 11(21):R872-874). Cdc14 activates protein kinase cdc15 to promote mitotic exit in budding yeast (Jaspersen S. L. and Morgan D. O., (2000) Curr Biol 10(10):615-618).

So far, it has not been described that the proteins of the invention and homologous proteins are involved in the regulation of energy homeostasis and body-weight regulation and related disorders, and thus, no functions in metabolic diseases and dysfunctions and other diseases as listed above have been discussed.

The function of Eip75B in metabolic disorders is further validated by data obtained from an additional screen. For example, an additional screen using Drosophila mutants with modifications of the eye phenotype identified a modification of UCP activity by Eip75B, thereby leading to an altered mitochondrial activity. These findings suggest the presence of similar activities of these described homologous proteins in humans that provides insight into diagnosis, treatment, and prognosis of metabolic disorders.

In this invention we refer to the proteins encoded by Drosophila Eip75B, Cdk4, or CG7134 genes and homologous orthologs, preferably human and

mice, homologous polypeptides or proteins or sequences encoding those proteins as proteins of the invention.

The present invention discloses that Eip75B, Cdk4, or CG7134 homologous proteins are regulating the energy homeostasis and fat metabolism especially the metabolism and storage of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides of the invention. The invention also relates to the use of these polynucleotides, polypeptides and effectors thereof in the diagnosis, study, prevention, and treatment of metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones.

Eip75B, Cdk4, or CG7134 homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids encoding the human Eip75B, Cdk4, or CG7134 homologs (in particular the human isoforms of NR1D1, NR1D2, CDK4, CDK6, CDC14A, CDC14B, MGC26484, or human protein similar to CDC14B isoform 3).

The invention particularly relates to a nucleic acid molecule encoding a polypeptide contributing to regulating the energy homeostasis and the metabolism of triglycerides, wherein said nucleic acid molecule comprises

(a) the nucleotide sequence of Drosophila Eip75B, Cdk4, or CG7134, human Eip75B, Cdk4, or CG7134 homologs (in particular the human isoforms of NR1D1, NR1D2, CDK4, CDK6, CDC14A, CDC14B, MGC26484, or human protein similar to CDC14B isoform 3), and/or a sequence complementary thereto,

(b) a nucleotide sequence which hybridizes at 65°C in a solution containing 0.2 x SSC and 0.1% SDS to a sequence of (a),

(c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,

(d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the amino acid sequences of the Eip75B, Cdk4, or CG7134 protein, preferably of the human Eip75B, Cdk4, or CG7134 homologs (in particular the human isoforms of NR1D1, NR1D2, CDK4, CDK6, CDC14A, CDC14B, MGC26484, or human protein similar to CDC14B isoform 3),

(e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or

(f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of at least 15 bases, preferably at least 20 bases, more preferably at least 25 bases and most preferably at least 50 bases.

The present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity, fragments of said genes, polypeptides encoded by said genes or fragments thereof, and effectors e.g. antibodies, biologically active nucleic acids, such as antisense molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

The ability to manipulate and screen the genomes of model organisms such as the fly Drosophila melanogaster provides a powerful tool to analyze biological and biochemical processes that have direct relevance to more complex vertebrate organisms due to significant evolutionary conservation of genes, cellular processes, and pathways (see, for example, Adams M. D. et al., (2000) Science 287: 2185-2195). Identification of novel gene

functions in model organisms can directly contribute to the elucidation of correlative pathways in mammals (humans) and of methods of modulating them. A correlation between a pathology model (such as changes in triglyceride levels as indication for metabolic syndrome including obesity) and the modified expression of a fly gene can identify the association of the human ortholog with the particular human disease.

In one embodiment, a forward genetic screen is performed in fly displaying a mutant phenotype due to misexpression of a known gene (see, Johnston Nat Rev Genet 3: 176-188 (2002); Rorth P., (1996) Proc Natl Acad Sci U S A 93: 12418-12422). In this invention, we have used a genetic screen to identify mutations of Eip75B, Cdk4, or CG7134 homologous genes that cause changes in the body weight which is reflected by a significant change of triglyceride levels. Triglycerides levels reflect the status of energy storage in cells and are significantly increased in obese patients. One resource for screening was a Drosophila melanogaster stock collection of EP-lines. The P-vector of this collection has Gal4-UAS-binding sites fused to a basal promoter that can transcribe adjacent genomic Drosophila sequences upon binding of Gal4 to UAS-sites (Brand & Perrimon (1993) Development 118:401-415; Rorth P., (1996) Proc Natl Acad Sci U S A 93:12418-12422). This enables the EP-line collection for overexpression of endogenous flanking gene sequences. In addition, without activation of the UAS-sites, integration of the EP-element into the gene is likely to cause a reduction of gene activity, and allows determining its function by evaluating the loss-of-function phenotype.

To isolate genes with a function in energy homeostasis, several thousand EP-lines were tested for their triglyceride content after a prolonged feeding period (see Examples for more detail). Lines with significantly changed triglyceride content were selected as positive candidates for further analysis. The change of triglyceride content due to the loss of a gene function suggests gene activities in energy homeostasis in a dose

dependent manner that control the amount of energy stored as triglycerides.

In this invention, the content of triglycerides of a pool of flies with the same genotype was analyzed after feeding for six days using a triglyceride assay. Male flies homozygous for the integration of vectors for Drosophila lines HD-EP(2)21120, or HD-EP(2)20271, or heterozygous for the integration of vectors for Drosophila line HD-EP(3)30293 were analyzed in assays measuring the triglyceride contents of these flies, illustrated in more detail in the EXAMPLES section. The results of the triglyceride content analysis are shown in FIGURES 1, 4, and 7.

An additional screen using Drosophila mutants with modifications of the eye phenotype identified a modification of UCP activity by Eip75B, thereby leading to an altered mitochondrial activity.

Genomic DNA sequences were isolated that are localized to the EP vector (herein HD-EP(3)30293, HD-EP(2)21120, OR HD-EP(2)20271) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly; see also FlyBase (1999) Nucleic Acids Research 27:85-88) were screened thereby identifying the integration site of the vectors, and the corresponding gene, described in more detail in the EXAMPLES section. The molecular organization of the gene is shown in FIGURES 2, 5, and 8.

In one embodiment of the invention, we clearly show that mammalian NR1D1 is expressed in varieties of mammalian (mouse) tissues, with highest levels of expression in metabolic active tissue such as brown adipose tissue (BAT) (see FIGURE 10A). In addition, expression of NR1D1 is strongly induced during the in vitro differentiation of 3T3-L1 (FIGURE 10B).

The invention also encompasses polynucleotides that encode the proteins of the invention and homologous proteins. Accordingly, any nucleic acid sequence, which encodes the amino acid sequences of the proteins of the invention and homologous proteins, can be used to generate recombinant molecules that express the proteins of the invention and homologous proteins. In a particular embodiment, the invention encompasses a nucleic acid encoding Drosophila Eip75B, Cdk4, or CG7134 or human Eip75B, Cdk4, or CG7134 homologs; referred to herein as the proteins of the invention. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl & Berger (1987: Methods Enzymol. 152:399-407) and Kimmel (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h in 0.2 x SSC and 0.1% SDS at 65°C a positive hybridization signal is observed. Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides or peptide mimetics having a length of at least 4, preferably at least 6 and up to 50 amino acids.

Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention and homologous proteins. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structures or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with the others, one or more times in a given sequence.

The nucleic acid sequences encoding the proteins of the invention and homologous proteins may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements.

In order to express a biologically active protein, the nucleotide sequences encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding

sequence. Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding the proteins and the appropriate transcriptional and translational control elements. Regulatory elements include for example a promoter, an initiation codon, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the Cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin promoter (see, Soria et al., 2000, Diabetes 49:157), SOX2 gene promotor (see Li et al., 1998, Curr. Biol. 8:971-4), Msi-1 promotor (see Sakakibara et al., 1997, J. Neuroscience 17:8300-8312), alpha-cardia myosin heavy chain promotor or human atrial natriuretic factor promotor (Klug et al., 1996, J. clin. Invest 98:216-24; Wu et al., 1989, J. Biol. Chem. 264:6472-79) or (iii) inducible promoters such as the tetracycline inducible system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention and homologous proteins may be ligated to a heterologous sequence to encode a fusion protein.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include, but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors;

yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus, adenovirus, adeno-associated virus, lentiverus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.

The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants containing DNA or RNA encoding the corresponding protein. As used herein 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide sequences include oligo-labeling, nick translation, end-labeling of labeled RNA probes, PCR amplification using a labeled nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

The presence of proteins of the invention in a sample can be determined by immunological methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for

determining protein activity are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

The nucleic acids encoding the proteins of the invention can be used to generate transgenic animal or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination, where the normal locus of the genes encoding the proteins of the invention is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retrovirusses and other animal virusses, YACs, and the like. The modified cells or animal are useful in the study of the function and regulation of the proteins of the invention. For example, a series of small deletions and/or substitutions may be made in the genes that encode the proteins of the invention to determine the role of particular domains of the protein, functions in pancreatic differentiation, etc.

Specific constructs of interest include anti-sense molecules, which will block the expression of the proteins of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z, may be introduced in the locus of the genes of the invention, where

upregulation of expression of the genes of the invention will result in an easily detected change in phenotype.

One may also provide for expression of the genes of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the proteins of the invention in cells in which they are not normally produced, one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and/or negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in presence of leukemia inhibiting factor (LIF).

When ES or embryonic cells or somatic pluripotent stem cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine

horn of pseudopregnant females. Females are then allowed to go to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected. The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, etc.

Diagnostics and Therapeutics

The data disclosed in this invention show that the nucleic acids and proteins of the invention and effector molecules thereof are useful in diagnostic and therapeutic applications implicated, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones.. Hence, diagnostic and therapeutic uses for the proteins of the invention nucleic acids and proteins of the invention are, for example but not limited to, the following: (i) protein therapeutic, (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic and/or prognostic marker, (v) gene therapy (gene delivery/gene ablation), (vi) research tools, and (vii) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues).

The nucleic acids and proteins of the invention and effectors thereof are useful in diagnostic and therapeutic applications implicated in various

applications as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, but not limited to, in metabolic disorders as described above.

The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that bind immunospecifically to the substances of the invention may be used in therapeutic or diagnostic methods.

For example, in one aspect, antibodies, which are specific for the proteins of the invention and homologous proteins, may be used directly as an effector, e.g. an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the protein. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. It is preferred that the peptides, fragments or oligopeptides used to induce antibodies to the protein have an

amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

Monoclonal antibodies to the proteins may be prepared using any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Köhler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R. J. et al. Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S. P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of 'chimeric antibodies', the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M. S. et al (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for the proteins of the invention and homologous proteins. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3). Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not

limited to, the $F(ab')_2$ fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reacive to two non-interfering protein epitopes are preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides of the invention or fragments thereof or nucleic acid effector molecules such as antisense molecules,aptamers, RNAi molecules, or ribozymes may be used for therapeutic purposes. In one aspect, aptamers, i.e. nucleic acid molceules, which are capable of binding to a Eip75B, Cdk4 or CG7134 homologous protein and of modulating its activity may be generated by a screening and selection method involving the use of combinatorial nucleic acid libraries. In a further aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding the proteins of the invention and homologous proteins. Thus, antisense molecules may be used to modulate protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the

coding or control regions of sequences encoding the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules complementary to the polynucleotides of the genes encoding the proteins of the invention and homologous proteins. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention and homologous proteins can be turned off by transforming a cell or tissue with expression vectors, which express high levels of polynucleotides that encode the proteins of the invention and homologous proteins or fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, e.g. DNA, RNA or nucleic acid analogues such as PNA, to the control regions of the genes encoding the proteins of the invention and homologous proteins, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it cause inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J. E. et al. (1994) In; Huber, B. E. and B. I. Carr, Molecular and Immunologic

Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples, which may be used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention and homologous proteins. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Nucleic acid effector molecules, e.g. an antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications

include, but are not limited to, the addition of flanking sequences at the 5prime and/or 3prime ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of the nucleic acids and the proteins of the invention and homologous nucleic acids or proteins, antibodies to the proteins of the invention and homologous proteins, mimetics, agonists, antagonists or inhibitors of the proteins of the invention and homologous proteins or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline,

buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones. The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compounds, the therapeutically effective does can be estimated initially either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active ingredient, for example the nucleid acids or the proteins of the invention and homologous proteins or nucleic acids or fragments thereof, antibodies of the proteins of the invention and homologous proteins, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g.,

ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage from employed, sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation. Normal dosage amounts may vary from 0.1 to 100,000 microg, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or associated with over- or underexpression of the proteins of the invention and homologous proteins or in assays to monitor patients being treated

with the proteins of the invention and homologous proteins, or effectors thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include methods which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene expression. Normal or standard values for gene expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of protein expressed in control and disease samples e.g. from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides specific for the proteins of the invention and homologous proteins may be used for diagnostic purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

In one aspect, hybridization with probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding the proteins of the invention and homologous proteins or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be DNA or RNA and are preferably derived from the nucleotide sequence of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be used for the diagnosis of conditions or diseases, which are associated with the expression of the proteins. Examples of such conditions or diseases include, but are not limited to, pancreatic diseases and disorders, including diabetes. Polynucleotide sequences specific for the proteins of the invention and homologous proteins may also be used to monitor the progress of patients receiving treatment for pancreatic diseases and disorders, including diabetes. The polynucleotide sequences may be used qualitative or quantitative assays, e.g. in Southern or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered gene expression.

In a particular aspect, the nucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be useful in assays that detect activation or induction of various metabolic diseases or dysfunctions, for example, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia

(dyslipidemia), and gallstones. The nucleotide sequences may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding the proteins of the invention and homologous proteins in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disease associated with expression of the proteins of the invention and homologous proteins, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a fragment thereof, which is specific for the nucleic acids encoding the proteins of the invention and homologous nucleic acids, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the normal patient. The

results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to metabolic diseases such as described above the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding the proteins of the invention and homologous proteins may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5prime.fwdarw.3prime) and another with antisense (3prime.rarw.5prime), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

In another embodiment of the invention, the nucleic acid sequences may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C. M. (1993)

Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154. FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or predisposition to a specific disease, may help to delimit the region of DNA associated with that genetic disease.

The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. An analysis of polymorphisms, e.g. single nucleotide polymorphisms may be carried out. Further, in situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

In another embodiment of the invention, the proteins of the invention, their catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds in any of a variety of drug screening techniques. One can identify effectors, e.g. receptors, enzymes, proteins, ligands, or substrates that bind to, modulate or mimic the action of one or more of the proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellulary. The formation of binding complexes, between the proteins of the invention and the agent tested, may be measured. Agents can also be identified, which either directly or indirectly influence the activity of the protein of the invention. Agents could also, either directly or indirectly, influence the activity of the proteins of the invention. For example the phosphatase activity of the proteins of the invention could be measured in vitro by using recombinantly expressed and purified Eip75B, Cdk4, or CG7134 or fragments thereof by making use of artificial substrates well known in the art, i.e. but not exclusively DiFMUP (Molecular Probes, Eugene, Oregon), which are converted to fluorophores or chromophores upon dephosphorylation. Alternatively, the dephosphorylation of physiological substrates of the phosphatases could be measured by making use of any of the well known screening technologies suitable for the detection of the phosphorylation status of their physiological substrates. For example, but not exclusively, the phosphorylation status of peptides derived from their physiological substrates can be monitored by binding of phospho-side specific antibodies resulting in an increase of the polarization of the complex.

In addition activity of Eip75B, Cdk4, or CG7134 against its physiological substrate(s) or derivatives thereof could be measured in cell-based assays. Agents may also interfere with posttranslational modifications of the protein, such as phosphorylation and dephosphorylation, farnesylation, palmitoylation, acetylation, alkylation, ubiquitination, proteolytic

processing, subcellular localization and degradation. Moreover, agents could influence the dimerization or oligomerization of the proteins of the invention or, in a heterologous manner, of the proteins of the invention with other proteins, for example, but not exclusively, docking proteins, enzymes, receptors, or translation factors. Agents could also act on the physical interaction of the proteins of this invention with other proteins, which are required for protein function, for example, but not exclusively, their downstream signaling.

Methods for determining protein-protein Interaction are well known in the art. For example binding of a fluorescently labeled peptide derived from the interacting protein to the protein of the Invention, or vice versa, could be detected by a change in polarisation. In case that both binding partners, which can be either the full length proteins as well as one binding partner as the full length protein and the other just represented as a peptide are fluorescently labeled, binding could be detected by fluorescence energy transfer (FRET) from one fluorophore to the other. In addition, a variety of commercially available assay principles suitable for detection of protein-protein Interaction are well known In the art, for example but not exclusively AlphaScreen (PerkinElmer) or Scintillation Proximity Assays (SPA) by Amersham. Alternatively, the interaction of the proteins of the invention with cellular proteins could be the basis for a cell-based screening assay, in which both proteins are fluorescently labeled and interaction of both proteins is detected by analysing cotranslocation of both proteins with a cellular imaging reader, as has been developed for example, but not exclusively, by Cellomics or EvotecOAI. In all cases the two or more binding partners can be different proteins with one being the protein of the invention, or in case of dimerization and/or oligomerization the protein of the invention itself. Proteins of the invention, for which one target mechanism of interest, but not the only one, would be such protein/protein interactions are Eip75B, Cdk4, or CG7134.

Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups.

Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal.

Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with the proteins or fragments thereof, and washed. Bound proteins are then detected by methods well known in the art. Purified proteins can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support. In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound for binding the protein. In this manner, the antibodies can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the protein.

Finally, the invention also relates to a kit comprising at least one of

(a) an Eip75B, Cdk4, or CG7134 nucleic acid molecule or a fragment thereof;

(b) an Eip75B, Cdk4, or CG7134 amino acid molecule or a fragment or an isoform thereof;

(c) a vector comprising the nucleic acid of (a);

(d) a host cell comprising the nucleic acid of (a) or the vector of (b);

(e) a polypeptide encoded by the nucleic acid of (a);

(f) a fusion polypeptide encoded by the nucleic acid of (a);

(g) an antibody, an aptamer or another receptor against the nucleic acid of (a) or the polypeptide of (d) or (e) and

(g) an anti-sense oligonucleotide of the nucleic acid of (a).

The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

The Figures show:

FIGURE 1 shows the content of energy storage triglycerides of Drosophila Eip75B (GadFly Accession Number CG8127) mutants. Shown is the change of triglyceride content of HD-EP(3)30293 flies caused by integration of the P-vector into the annotated transcription unit (column 2) in comparison to controls containing all fly lines of the proprietary EP collection ('EP-control)', column 1)

FIGURE 2 shows the molecular organization of the mutated Eip75B (GadFly Accession Number CG8127) gene locus.

FIGURE 3 shows the nucleic acid and amino acid sequences of the human nuclear receptor subfamily 1, group D (NR1D).
Figure 3A shows the nucleic acid sequence of human NR1D1 (SEQ ID NO: 1)
Figure 3B shows the amino acid sequence (one-letter code) of human NR1D1 (SEQ ID NO: 2).
Figure 3C shows the nucleic acid sequence of human NR1D2 (SEQ ID NO: 3)
Figure 3D shows the amino acid sequence (one-letter code) of human NR1D2 (SEQ ID NO: 4).

FIGURE 4 shows the content of energy storage triglycerides of Drosophila Cdk4 (GadFly Accession Number CG5072) mutants. Shown is the change of triglyceride content of HD-EP(2)21120 flies caused by integration of the P-vector into the annotated transcription unit (column 2) in comparison to

controls containing all fly lines of the proprietary EP collection ('EP-control)', column 1)

FIGURE 5 shows the molecular organization of the mutated Cdk4 (GadFly Accession Number CG5072) gene locus.

FIGURE 6 shows the nucleic acid and amino acid sequences of the human cyclin-dependent kinases (CDK).
Figure 6A shows the nucleic acid sequence of human CDK6 (SEQ ID NO: 5)
Figure 6B shows the amino acid sequence (one-letter code) of human CDK6 (SEQ ID NO: 6).
Figure 6C shows the nucleic acid sequence of human CDK4 (SEQ ID NO: 7)
Figure 6D shows the amino acid sequence (one-letter code) of human CDK4 (SEQ ID NO: 8).

FIGURE 7 shows the content of triglyceride of Drosophila CG7134 (GadFly Accession Number) mutants. Shown is the change of triglyceride content of HD-EP(2)20271 flies caused by integration of the P-vector into the annotated transcription unit (column 3) in comparison to controls containing more than 2000 fly lines of the proprietary EP collection ('HD-control (TG)', column 1) and wildtype controls determined in more than 80 independent assays (referred to as 'WT-control (TG)' column 2).

FIGURE 8 shows the molecular organization of the mutated CG7134 (GadFly Accession Number) gene locus.

FIGURE 9 shows the nucleic acid and amino acid sequences of the human cell division cycle 14 (CDC14) homologs A, B, and MGC26484
Figure 9A shows the nucleic acid sequence of human CDC14 homolog A, transcript variant 1 (SEQ ID NO: 9)

Figure 9B shows the amino acid sequence (one-letter code) of human CDC14 homolog A, isoform 1 (SEQ ID NO: 10)

Figure 9C shows the nucleic acid sequence of human CDC14 homolog A, transcript variant 2 (SEQ ID NO: 11)

Figure 9D shows the amino acid sequence (one-letter code) of human CDC14 homolog A, isoform 2 (SEQ ID NO: 12)

Figure 9E shows the nucleic acid sequence of human CDC14 homolog A, transcript variant 3 (SEQ ID NO: 13)

Figure 9F shows the amino acid sequence (one-letter code) of human CDC14 homolog A, isoform 3 (SEQ ID NO: 14)

Figure 9G shows the nucleic acid sequence of human CDC14 homolog B, transcript variant 1 (SEQ ID NO: 15)

Figure 9H shows the amino acid sequence (one-letter code) of human CDC14 homolog B, isoform 1 (SEQ ID NO: 16)

Figure 9I shows the nucleic acid sequence of human CDC14 homolog B, transcript variant 2 (SEQ ID NO: 17)

Figure 9J shows the amino acid sequence (one-letter code) of human CDC14 homolog B, isoform 2 (SEQ ID NO: 18)

Figure 9K shows the nucleic acid sequence of human CDC14 homolog B, transcript variant 3 (SEQ ID NO: 19)

Figure 9L shows the amino acid sequence (one-letter code) of human CDC14 homolog B, isoform 3 (SEQ ID NO: 20)

Figure 9M shows the nucleic acid sequence of human hypothetical protein MGC26484 (MGC26484) (SEQ ID NO: 21)

Figure 9N shows the amino acid sequence (one-letter code) of human hypothetical protein MGC26484 (MGC26484) (SEQ ID NO: 22)

Figure 9O shows the nucleic acid sequence of human protein similar to CDC14 homolog B, isoform 3 (SEQ ID NO: 23)

Figure 9P shows the amino acid sequence (one-letter code) of human protein similar to CDC14 homolog B, isoform 3 (SEQ ID NO: 24)

Figure 10 shows the expression of Nr1D1 in different mammalian models

Figure 10A. Real-time PCR analysis of NR1D1 expression in wildtype mouse tissues. The relative RNA-expression is shown on the Y-axis, the tissues tested are given on the X-axis. (WAT = white adipose tissue; BAT = brown adipose tissue)

Figure 10B shows real-time PCR analysis of NR1D1 expression during differentiation of mammalian fibroblast (3T3-L1) cells from pre-adipocytes to mature adipocytes.

The examples illustrate the invention:

Example 1: Measurement of triglyceride content

Mutant flies are obtained from a fly mutation stock collection. The flies are grown under standard conditions known to those skilled in the art. In the course of the experiment, additional feedings with bakers yeast (Saccharomyces cerevisiae) are provided for the EP-lines HD-EP(3)30293, HD-EP(2)21120, OR HD-EP(2)20271. The average change of triglyceride content of Drosophila containing the EP-vector as homozygous or heterozygous viable integration was investigated in comparison to control flies, respectively (see FIGURES 1, 4, and 7). For determination of triglyceride content, flies were incubated for 5 min at 90°C in an aqueous buffer using a waterbath, followed by hot extraction. After another 5 min incubation at 90°C and mild centrifugation, the triglyceride content of the flies extract was determined using Sigma Triglyceride (INT 336-10 or -20) assay by measuring changes in the optical density according to the manufacturer's protocol. As a reference the protein content of the same extract was measured using BIO-RAD DC Protein Assay according to the manufacturer's protocol. These experiments and assays were repeated several times.

The average triglyceride level (microg triglyceride/microg protein) of all flies of the EP collection (referred to as 'EP-control') is shown as 100% in the first column in FIGURES 1 and 4. The average triglyceride level (microg

triglyceride/microg protein) of 2108 fly lines of the proprietary EP-collection (referred to as 'HD-control (TG)') is shown as 100% in the first column in FIGURE 7. The average triglyceride level (microg triglyceride/microg protein) of Drosophila wildtype strain Oregon R flies determined in 84 independent assays (referred to as 'WT-control (TG)') is shown as 102% in the second column in FIGURE 7. Standard deviations of the measurements are shown as thin bars.

HD-EP(3)30293 heterozygous flies show constantly a higher triglyceride content (microg triglyceride/microg protein) than the controls (column 2 in FIGURE 1, 'HD-EP30293 / TM3'). Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage triglycerides.

HD-EP(2)21120 homozygous flies show constantly a higher triglyceride content (microg triglyceride/microg protein) than the controls (column 2 in FIGURE 4, 'HD-EP21120'). Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage triglycerides.

HD-EP(2)20271 homozygous flies show constantly a higher triglyceride content (microg triglyceride/microg protein) than the controls (column 3 in FIGURE 7, 'HD-20271 (TG)'). Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage triglycerides.

Example 2: Identification of the genes

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(3)30293) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous lethal

integration site of the HD-EP(3)30293 vector into the third exon of the transcript variant CG8127-RA of the gene Eip75B in antisense orientation. FIGURE 2 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector of HD-EP(3)30293 is at gene locus 3L, 75A8-B2. In FIGURE 2, genomic DNA sequence is represented by the assembly as a black arrow in middle of the figure that includes the integration site of HD-EP(3)30293. Ticks represent the length in basepairs of the genomic DNA (10000 base pairs per tick). Black bars in the lower half of the figure, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project). Predicted exons of the Drosophila cDNA of the gene Eip75B (GadFly Accession Number CG8127) are shown as black bars and predicted introns as slim black lines in the lower half of the figure and are labeled. The integation site of HD-EP(3)30293 is indicated with a black triangle within the first exon of the Eip75B predicted cDNA transcript variant RA. Therefore, expression of the cDNA encoding Eip75B could be affected by integration of the vector of line HD-EP(3)30293, leading to a change in the amount of energy storage triglycerides.

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(2)21120) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(2)21120 vector into an intron of the cDNA of the gene Cdk4 in sense orientation. FIGURE 5 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector of HD-EP(2)21120 is at gene locus 2R, 53C9 (according to Flybase and Gadfly release 3). In FIGURE 5, genomic DNA sequence is represented by the assembly as a black arrow in the lower half of the figure that includes the integration site of HD-EP(2)21120. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs per tick). Dark grey bars in the middle of the figure, linked by dark grey lines

represent cDNAs of the predicted gene (as predicted by the Berkeley Drosophila Genome Project). Predicted exons of the Drosophila cDNA of the gene Cdk4 (GadFly Accession Number CG5072) are shown as black bars and predicted introns as slim black lines in the middle of the figure and are labeled. The integation site of HD-EP(2)21120 is indicated with a black triangle within an intron of the Cdk4 predicted cDNA transcript variants. Therefore, expression of the cDNA encoding Cdk4 could be affected by integration of the vector of line HD-EP(2)21120, leading to a change in the amount of energy storage triglycerides.

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(2)20271) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(2)20271 vector into the first intron, 823 base pairs 3prime of the first exon of the cDNA of the gene CG7134 in antisense orientation. FIGURE 8 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector of HD-EP(2)20271 is at gene locus 2L, 28C1 (according to Flybase), or 28C4-5 (according to Gadfly release 3). In FIGURE 8, genomic DNA sequence is represented by the assembly as a dotted grey line in middle of the figure that includes the integration site of HD-EP(2)20271. Numbers represent coordinates of the genomic DNA (starting at position 7793500 on chromosome 2L, ending at position 7804500 on chromosome 2L). The insertion site of the P-element in Drosophila line HD-EP(2)20271 is shown as bar in the "P Elements +" line and is labeled. Dark grey bars on the "cDNA +" and the "cDNA -" lines, partly linked by light grey bars, represent the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly and by Magpie). Predicted exons are shown as dark grey bars, predicted introns are shown as light grey bars. The gene CG7134 is labeled. Transcribed DNA sequences (ESTs) are shown as grey bars in the "EST +" and "EST -" lines. Therefore, expression of the cDNA

encoding CG7134 could be affected by integration of the vector of line HD-EP(2)20271, leading to a change in the amount of energy storage triglycerides.

Table 1 is summarizing the data of our molecular analysis of the Drosophila protein identified in this invention as being involved in the regulation of the metabolism.

Table 1. Molecular analysis of Drosophila Eip75B, Cdk4, and CG7134

Analysis	Genetic interaction
Eip75B	with *Kr* and *ph-p*
Cdk4	Retinoblastoma-family protein (Rbf, DNA binding), Cyclin D (CycD), gigas (gig), Tsc1, patched (ptc, receptor signaling) (Flybase)
CG7134	not described (Flybase)
Analysis	**Protein**
Eip75B	ligand-dependent nuclear receptor (Flybase)
Cdk4	cyclin-dependent protein kinase (Flybase)
CG7134	protein tyrosine/serine/threonine phosphatase (Flybase)
Analysis	**Protein domains**
Eip75B	Ligand-binding domain of nuclear hormone receptor, C4-type steroid receptor zink finger, Steroid hormone receptor, thyroid hormone receptor, Nuclear receptor ROR family (1F nuclear receptor), Glucocorticoid receptor like (DNA-binding domain) (Flybase)
Cdk4	Eukaryotic protein kinase, Serine/Threonine protein kinase family active site, Protein kinase-like (PK-like) (Flybase)
CG7134	Dual specificity protein phosphatase, Tyrosine specific protein phosphatase and dual specificity protein phosphatase family, (Phosphotyrosine protein) phosphatases II (Flybase)
Analysis	**InterPro analysis**
Eip75B	-
Cdk4	Eukaryotic protein kinase (IPR000719), Tyrosine protein kinase (IPR001245), Serine/Threonine protein kinase (IPR002290)
CG7134	Tyrosine specific protein phosphatase (IPR000242), Dual specificity protein phosphatase (IPR000340), Tyrosine specific protein phosphatase and dual specificity protein phosphatase (IPR000387)
Analysis	**Locus**
Eip75B	3L 75A8-B2 (GadFly)
Cdk4	2R, 53C9 (Flybase); 2R, 53C9 (Gadfly release 3)
CG7134	2L, 28C1 (Flybase); 2L 28C4-5 (Gadfly release 3)
Analysis	**Ests**
Eip75B	
Cdk4	several including LD31205 (Gadfly release 3)
CG7134	several
Analysis	**cDNA**
Eip75B	

Cdk4	AA246773 (605 bp mRNA, 2001), AA950999 (759 bp mRNA, 2001), AW943070 (555 bp mRNA, 2001), AY060397 (2027 bp mRNA, 2001; protein:AAL25436), X99510 (1132 bp mRNA, 1996; protein:CAA67860) (Flybase)
CG7134	not described (Flybase)
Analysis	**genomic DNA**
Eip75B	
Cdk4	AC005647 (59880 bp DNA, 1999), AE003806 (268219 bp DNA, 2000; protein:AAF57980; protein:AAF57981; protein:AAM68505) (Flybase)
CG7134	AE003618 (268667 bp DNA, 2000; protein:AAF52562) (Flybase)
Analysis	**NCBI locus ID**
Eip75B	39999, Aliases: 57B, E75, E75A, E75B, E75C, dE75, E75-C, Eip75, NR1D3, CG8127, DmE75A, DmE75B, CT24290, EP1121b, l(3)j3A6, l(3)j5E1, l(3)07041, l(3)j11A6, l(3)j12E8, l(3)neo25; RefSeq: NM_079409; Nucleotide: AE003522, AQ073338, AQ073339, AQ073765, BH146157, BH609996, Z83526, AA696061, AW941240, X15586, X51548, X51549; Protein: NP_524133, AAF49282, CAA33611, CAA35923, CAA35924
Cdk4	36854, Dm Cdk4, Cyclin-dependent kinase 4, 53C9; Aliases: 8-6, CDK4, Pk?7, cdk4, Pk53C, CDK4/6, CG5072, Cdk4/6, DmCdk4, cdk4/6, CT15896, CT16072, l(2)05428, l(2)s4639, l(2)k06503; RefSeq: NM_057848; Nucleotide: AC004287, AC005647, AE003806, AE003806, AQ024975, AQ025637, AQ025797, AQ073382, AQ073523, AA246773, AA696227, AA950849, AA978713, AW943070, AY060397, BI363050, BI372046, X99510; Protein: NP_477196, AAF57980, AAF57981, AAM68505, AAL25436, CAA67860
CG7134	34067, Dm CG7134, 28C1; Aliases: CDC14, CT22031; Nucleotide: AE003618; Protein: AAF52562
Analysis	**Drosophila mutations & mutants**
Eip75B	there are 60 recorded mutant alleles , of which at least 4 are available from the public stock centers (Flybase).
Cdk4	There are 12 recorded alleles: 4 in vitro constructs (1 available from the public stock centers), 7 classical mutants (3 available from the public stock centers) and 1 wild-type. (Flybase)
CG7134	not described (Flybase)
Analysis	**Phenotypic info**

Eip75B	not described (Flybase)
Cdk4	Mutations have been isolated which affect the embryonic cuticle, the maternal effect segment and the maternal effect head and are larval recessive lethal. (Flybase)
CG7134	not described (Flybase)

Example 3: Identification of the human Eip75B, Cdk4, or CG7134 homologous proteins

Eip75B, Cdk4, or CG7134 homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids comprising Drosophila Eip75B, Cdk4, or CG7134 or human Eip75B, Cdk4, or CG7134 and homologs. Sequences homologous to Drosophila Eip75B, Cdk4, or CG7134 were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI) (see, Altschul et al., 1997, Nucleic Acids Res. 25:3389-3402). Table 2 shows the best human homologs of the Drosophila Eip75B, Cdk4, or CG7134 genes.

The term "polynucleotide comprising the nucleotide sequence as shown in GenBank Accession number" relates to the expressible gene of the nucleotide sequences deposited under the corresponding GenBank Accession number. The term "GenBank Accession number" relates to NCBI GenBank database entries (Ref.: Benson et al., Nucleic Acids Res. 28 (2000) 15-18). The term "IPI Accession Number" relates to ENSEMBL International Protein Index entries (http://www.ensembl.org/IPI/; Hubbard T. et al., (2002) Nucleic Acids Research 30 (1): 38-41)

Table 2. Human homologous proteins to Drosophila Eip75B, Cdk4, or CG7134 protein

I. Eip75B

NCBI (National Center for Biotechnology Information) human locus identification (ID): 9572, Hs NR1D1, nuclear receptor subfamily 1, group D, member 1, Position 17q11.2

Aliases: EAR1, hRev, EAR-1, THRA1, THRAL, REV-ERBAALPHA

OMIM: 602408

RefSeq[R]: GenBank Accession Number NM_021724

Nucleotide: GenBank Accession Numbers X95536, M24898, M24900, M34340, X72631

Protein: GenBank Accession Numbers NP_068370, AAA52335, AAA52332, AAA36562, CAB53540

NCBI (National Center for Biotechnology Information) human locus identification (ID): 9975, Hs NR1D2, nuclear receptor subfamily 1, group D, member 2, Position 3

Aliases: RVR, BD73, HZF2, EAR-1R, Hs.37288

OMIM: 602304

RefSeq[R]: GenBank Accession Numbers NM_005126, XM_171049

Nucleotide: GenBank Accession Numbers BC015929, D16815, L31785

Protein: GenBank Accession Numbers NP_005117, XP_171049, AAH15929, BAA20088, AAA65937

* * * * * * * * * * * * * *

II. Cdk4

NCBI (National Center for Biotechnology Information) human locus identification (ID): 1021, Hs CDK6, cyclin-dependent kinase 6, 7q21-q22

Aliases: PLSTIRE

OMIM: 603368

RefSeq[R]: GenBank Accession Number NM_001259

Nucleotide: GenBank Accession Numbers AC000065, AC004011, AC004128, X66365

Protein: GenBank Accession Numbers NP_001250, AAB46347, AAB96868, AAC01770, CAA47008

NCBI (National Center for Biotechnology Information) human locus identification (ID): 1019, Hs CDK4, cyclin-dependent kinase 4, 12q14

Aliases: CMM3, PSK-J3, MGC14458

OMIM: 123829

RefSeq[R]: GenBank Accession Numbers NM_000075, NM_052984

Nucleotide: GenBank Accession Numbers U37022, U81031, BC003644, BC005864, BC007968, BC010153, BC015669, M14505, U79269, Z48970

Protein: GenBank Accession Numbers NP_000066, NP_443710, AAC50506, AAC39521, AAH03644, AAH05864, AAH10153, AAH15669, AAA35673, AAB50213, CAA88834

III. CG7134

NCBI (National Center for Biotechnology Information) human locus identification (ID): 8556, Hs CDC14A, CDC14 cell division cycle 14 homolog A (S. cerevisiae), 1p21

Aliases: cdc14, hCDC14, Cdc14A1, Cdc14A2

OMIM: 603504

RefSeq[R]: GenBank Accession Numbers NM_003672, NM_033312, NM_033313

Nucleotide: GenBank Accession Numbers AF000367, AF064102, AF064103, AF122013

Protein: GenBank Accession Numbers NP_003663, NP_201569, NP_201570, AAB88277, AAC16659, AAC16660, AAD49217

NCBI (National Center for Biotechnology Information) human locus identification (ID): 8555, Hs CDC14B, CDC14 cell division cycle 14 homolog B (S. cerevisiae), 9q22.32

Aliases: CDC14B3, Cdc14B1, Cdc14B2, hCDC14B

OMIM: 603505

RefSeq[R]: GenBank Accession Numbers NM_003671, NM_033331, NM_033332

Nucleotide: GenBank Accession Numbers AF023158, AF064104, AF064105

Protein: GenBank Accession Numbers NP_003662, NP_201588, NP_201589, AAB88293, AAC16661, AAC16662

NCBI (National Center for Biotechnology Information) human locus identification (ID): 168448, Hs MGC26484, hypothetical protein MGC26484, 7p14.3

RefSeq: GenBank Accession Numbers NM_152627, XM_095105

Nucleotide: GenBank Accession Numbers AC006024, BC028690

Protein: GenBank Accession Numbers NP_689840, XP_095105, AAD15415, AAH28690

NCBI (National Center for Biotechnology Information) human locus identification (ID): similar to CDC14 homolog B, isoform 3; S. cerevisiae CDC14 homolog, gene B; CDC14 (cell division cycle 14, S. cerevisiae) homolog B, 7

RefSeq: GenBank Accession Number XM_171149

Protein: GenBank Accession Number XP_171149

The mouse homologous cDNAs encoding the polypeptides of the invention were identified as GenBank Accession Numbers NM_145434, XM_126627 (for the mouse homologs of NR1D1), NM_009870 (for the mouse homolog of Cdk4), NM_009873 (for the mouse homolog of Cdk6), and IPI Accession Numbers IPI00134094 and IPI00111965 (for the mouse homologs of CDC14A), and IPI Accession Numbers IPI00126617, IPI00126961, and IPI00124775 (for the mouse homologs of CDC14B).

Example 4: dUCPy modifier screen

Expression of Drosophila uncoupling protein dUCPy in a non-vital organ like the eye (Gal4 under control of the eye-specific promoter of the "eyeless" gene) results in flies with visibly damaged eyes. This easily visible eye phenotype is the basis of a genetic screen for gene products that can modify UCP activity.

Parts of the genomes of the strain with Gal4 expression in the eye and the strain carrying the pUAST-dUCPy construct were combined on one chromosome using genomic recombination. The resulting fly strain has eyes that are permanently damaged by dUCPy expression. Flies of this strain were crossed with flies of a large collection of mutagenized fly strains. In this mutant collection a special expression system (EP-element, Ref.: Rørth P, Proc Natl Acad Sci U S A 1996, 93(22):12418-22) is integrated randomly in different genomic loci. The yeast transcription factor Gal4 can bind to the EP-element and activate the transcription of endogenous genes close the integration site of the EP-element. The activation of the genes therefore occurs in the same cells (eye) that overexpress dUCPy. Since the mutant collection contains several thousand strains with different integration sites of the EP-element it is possible to test a large number of genes whether their expression interacts with dUCPy activity. In case a gene acts as an enhancer of UCP activity the eye defect will be worsened; a suppressor will ameliorate the defect.

Using this screen a gene with suppressing activity was discovered that was found to be the Eip75B gene in Drosophila.

Example 5: Expression of the polypeptides in mammalian (mouse) tissues
To analyse the expression of the polypeptides disclosed in this invention in mammalian tissues, several mouse strains (preferably mice strains C57Bl/6J, C57Bl/6 ob/ob and C57Bl/KS db/db which are standard model

systems in obesity and diabetes research) were purchased from Harlan Winkelmann (33178 Borchen, Germany) and maintained under constant temperature (preferably 22°C), 40 per cent humidity and a light / dark cycle of preferably 14 / 10 hours. The mice were fed a standard chow (for example, from ssniff Spezialitäten GmbH, order number ssniff M-Z V1126-000). For the fasting experiment ("fasted wild type mice"), wild type mice were starved for 48 h without food, but only water supplied ad libitum (see, for example, Schnetzler et al. J Clin Invest 1993 Jul;92(1):272-80, Mizuno et al. Proc Natl Acad Sci U S A 1996 Apr 16;93(8):3434-8). Animals were sacrificed at an age of 6 to 8 weeks. The animal tissues were isolated according to standard procedures known to those skilled in the art, snap frozen in liquid nitrogen and stored at -80°C until needed.

For analyzing the role of the proteins disclosed in this invention in the in vitro differentiation of different mammalian cell culture cells for the conversion of pre-adipocytes to adipocytes, mammalian fibroblast (3T3-L1) cells (e.g., Green & Kehinde, Cell 1: 113-116, 1974) were obtained from the American Tissue Culture Collection (ATCC, Hanassas, VA, USA; ATCC- CL 173). 3T3-L1 cells were maintained as fibroblasts and differentiated into adipocytes as described in the prior art (e.g., Qiu. et al., J. Biol. Chem. 276:11988-95, 2001; Slieker et al., BBRC 251: 225-9, 1998). In brief, cells were plated in DMEM/10% FCS (Invitrogen, Karlsruhe, Germany) at 50,000 cells/well in duplicates in 6-well plastic dishes and cultured in a humidified atmosphere of 5% CO2 at 37°C. At confluence (defined as day 0: d0) cells were transferred to serum-free (SF) medium, containing DMEM/HamF12 (3:1; Invitrogen), Fetuin (300 microg/ml; Sigma, Munich, Germany), Transferrin (2 microg/ml; Sigma), Pantothenate (17microM; Sigma), Biotin (1microM; Sigma), and EGF (0.8nM; Hoffmann-La Roche, Basel, Switzerland). Differentiation was induced by adding Dexamethasone (DEX; 1microM; Sigma), 3-Methyl-Isobutyl-1-Methylxanthine (MIX; 0.5mM; Sigma), and bovine

Insulin (5microg/ml; Invitrogen). Four days after confluence (d4), cells were kept in SF medium, containing bovine Insulin (5microg/ml) until differentiation was completed. At various time points of the differentiation procedure, beginning with day 0 (day of confluence) and day 2 (hormone addition; for example, dexamethason and 3-isobutyl-1-methylxanthin), up to 10 days of differentiation, suitable aliquots of cells were taken every two days.

RNA was isolated from mouse tissues or cell culture cells using Trizol Reagent (for example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in combination with an DNase-treatment according to the instructions of the manufacturers and as known to those skilled in the art. Total RNA was reverse transcribed (preferably using Superscript II RNaseH- Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the Taqman 2xPCR Master Mix (from Applied Biosystems, Weiterstadt, Germany; the Mix contains according to the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence Detection System (from Applied Biosystems, Weiterstadt, Germany).

The following prime/probe pairs were used for the TaqMan analysis:

For the amplification of mNR1D1
Mouse mNR1D1 forward primer (SEQ ID NO: 25): 5prime-CGG CTC AGC GTC ATA ATG AAG-3prime;
mouse mNR1D1 reverse primer (SEQ ID NO: 26): 5prime-AGG CCA GGT AGG CGG GTA-3prime;
Taqman probe (SEQ ID NO: 27): (5/6-FAM) CTG AAT GGT CTA CGC CAG GGC CC (5/6-TAMRA)

As shown in Figure 10A analysis of the expression of NR1D1 in mammalian (mouse) tissues revealed that NR1D1 is expressed in most mammalian tissues, showing highest level of expression in BAT, muscle, lung and heart. These results suggest an essential role for NR1D1 in most cell types.

As shown in Figure 10B, real time PCR (Taqman) analysis of the expression of the mNR1D1 revealed that NR1D1 is upregulated during 3T3-L1 differentiation. The observed 11 fold upregulation of this transcriptional regulator during adipogenesis suggests a critical role in adipocyte maturation, making it an interesting candidate gene for treating metabolic disorders.

Example 6: In vitro assays for the determination of triglyceride and glycogen storage

Obesity is known to be caused by different reasons such as non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure. For example, an increase in energy expenditure (and thus, lowering the body weight) would include the elevated utilization of both circulating and intracellular glucose and triglycerides, free or stored as glycogen or lipids as fuel for energy and/or heat production. In this invention, we therefore show the cellular level of triglycerides and glycogen in cells overexpressing the protein of the invention.

Retroviral infection of preadipocytes

Packaging cells were transfected with retroviral plasmids pLPCX carrying the mouse transgene encoding a protein of the invention and a selection marker using calcium phosphate procedure. Control cells were infected with pLPCX carrying no transgene. Briefly, exponentially growing packaging cells were seeded at a density of 350,000 cells per 6-well in 2 ml DMEM + 10 % FCS one day before transfection. 10 min before

transfection chloroquine was added directly to the overlying medium (25 microM end concentration). A 250 microliter transfection mix consisting of 5 micro plasmid-DNA (candidate:helper-virus in a 1:1 ratio) and 250 mM CaCl2 was prepared in a 15 ml plastic tube. The same volume of 2 x HBS (280 microM NaCl, 50 microM HEPES, 1.5 mM Na_2HPO_4, pH 7.06) was added and air bubbles were injected into the mixture for 15 sec. The transfection mix was added drop wise to the packaging cells, distributed and the cells were incubated at 37°C, 5 % CO_2 for 6 hours. The cells were washed with PBS and the medium was exchanged with 2 ml DMEM + 10 % CS per 6-well. One day after transfection the cells were washed again and incubated for 2 days of virus collection in 1 ml DMEM + 10 % CS per 6-well at 32°C, 5 % CO2. The supernatant was then filtered through a 0.45 μm cellulose acetate filter and polybrene (end concentration 8 μg/ml) was added. Mammalian fibroblast (3T3-L1) cells in a sub-confluent state were overlaid with the prepared virus containing medium. The infected cells were selected for 1 week with 2 μg/ml puromycin. Following selection the cells were checked for transgene expression by western blot and immunofluorescence. Over expressing cells were seeded for differentiation.

3T3-L1 cells were maintained as fibroblasts and differentiated into adipocytes as described in the prior art and supra. For analysing the role of the proteins disclosed in this invention in the in vitro assays for the determination of triglyceride storage, synthesis and transport were performed.

Preparation of cell lysates for analysis of metabolites

Starting at confluence (D0), cell media was changed every 48 hours. Cells and media were harvested 8 hours prior to media change as follows. Media was collected, and cells were washed twice in PBS prior to lyses in 600 microl HB-buffer (0.5% Polyoxyethylene 10 tridecylethan, 1 mM EDTA, 0.01M NaH_2PO_4, pH 7.4). After inactivation at 70°C for 5 minutes, cell lysates were prepared on Bio 101 systems lysing matrix B (0.1 mm silica

beads; Q-Biogene, Carlsbad, USA) by agitation for 2 x 45 seconds at a speed of 4.5 (Fastprep FP120, Bio 101 Thermosavant, Holbrock, USA). Supernatants of lysed cells were collected after centrifugation at 3000 rpm for 2 minutes, and stored in aliquots for later analysis at -80°C.

Changes in cellular triglyceride levels during adipogenesis

Cell lysates and media were simultaneously analysed in 96-well plates for total protein and triglyceride content using the Bio-Rad DC Protein assay reagent (Bio-Rad, Munich, Germany) according to the manufacturer's instructions and a modified enzymatic triglyceride kit (GPO-Trinder; Sigma) briefly final volumes of reagents were adjusted to the 96-well format as follows: 10 microl sample was incubated with 200 microl reagent A for 5 minutes at 37°C. After determination of glycerol (initial absorbance at 540 nm), 50 microl reagent B was added followed by another incubation for 5 minutes at 37°C (final absorbance at 540 nm). Glycerol and triglyceride concentrations were calculated using a glycerol standard set (Sigma) for the standard curve included in each assay.

Changes in cellular glycogen levels during adipogenesis

Cell lysates and media were simultaneously analysed in triplicates in 96-well plates for total protein and glycogen content using the Bio-Rad DC Protein assay reagent (Bio-Rad, Munich, Germany) according to the manufacturer's instructions and an enzymatic starch kit from Hoffmann-La Roche (Basel, Switzerland). 10-microL samples were incubated with 20-microliters amyloglucosidase solution for 15 minutes at 60°C to digest glycogen to glucose. The glucose is further metabolised with 100 microL distilled water and 100 microl of enzyme cofactor buffer and 12 microliters of enzyme buffer (hexokinase and glucose phosphate dehydrogenase). Background glucose levels are determined by subtracting values from a duplicate plate without the amyloglucosidase. Final absorbance is determined at 340 nm. HB-buffer as blank, and a standard curve of

glycogen (Hoffmann-La Roche) were included in each assay. Glycogen content in samples were calculated using a standard curve.

Synthesis of lipids during adipogenesis

During the terminal stage of adipogenesis (day 12) cells were analysed for their ability to metabolise lipids. A modified protocol to the method of Jensen et al (2000) for lipid synthesis was established. Cells were washed 3 times with PBS prior to serum starvation in Krebs-Ringer-Bicarbonate-Hepes buffer (KRBH; 134 nM NaCl, 3.5 mM KCl, 1.2 mM KH_2 PO_4, 0.5 mM $MgSO_4$, 1.5 mM $CaCl_2$, 5 mM $NaHCO_3$, 10 mM Hepes, pH 7.4), supplemented with 0.1% FCS for 2.5h at 37°C. For insulin-stimulated lipid synthesis, cells were incubated with 1 microM bovine insulin (Sigma; carrier: 0.005N HCl) for 45 min at 37°C. Basal lipid synthesis was determined with carrier only. 14C(U)-D-Glucose (NEN Life Sciences) in a final activity of 1microCi/Well/ml in the presence of 5 mM glucose was added for 30 min at 37°C. For the calculation of background radioactivity, 25 microM Cytochalasin B (Sigma) was used. All assays were performed in duplicate wells. To terminate the reaction, cells were washed 3 times with ice cold PBS, and lysed in 1 ml 0.1N NaOH. Protein concentration of each well was assessed using the standard Biuret method (Protein assay reagent; Bio-Rad). Total lipids were separated from aqueous phase after overnight extraction in Insta-Fluor scintillation cocktail (Packard Bioscience) followed by scintillation counting.

Transport and metabolism of free fatty acids during adipogenesis

During the terminal stage of adipogenesis (D12) cells were analysed for their ability to transport long chain fatty acid across the plasma membrane. A modified protocol to the method of Abumrad et al (1991) (Proc. Natl. Acad. Sci. USA, 1991: 88; 6008-12) for cellular transportation of fatty acid was established. In summary, cells were washed 3 times with PBS prior to serum starvation. This was followed by incubation in Krebs-Ringer-Bicarbonate-Hepes buffer (KRBH; 134 nM NaCl, 3.5 mM KCl,

1.2 mM KH_2PO_4, 0.5 mM $MgSO_4$, 1.5 mM $CaCl_2$, 5 mM $NaHCO_3$, 10 mM Hepes, pH 7.4), supplemented with 0.1% FCS for 2.5h at 37°C. Uptake of exogenous free fatty acids was initiated by the addition of isotopic media containing non radioactive oleate and (3H)oleate (NEN Life Sciences) complexed to serum albumin in a final activity of 1microCi/Well/ml in the presence of 5 mM glucose for 30min at room temperature (RT). For the calculation of passive diffusion (PD) in the absence of active transport (AT) across the plasma membrane 20mM of phloretin in glucose free media (Sigma) was added for 30 min at RT. All assays were performed in duplicate wells. To terminate the active transport 20mM of phloretin in glucose free media was added to the cells. Cells were lysed in 1 ml 0.1N NaOH and the protein concentration of each well were assessed using the standard Biuret method (Protein assay reagent; Bio-Rad). Esterified fatty acids were separated from free fatty acids using overnight extraction in Insta-Fluor scintillation cocktail (Packard Bioscience) followed by scintillation counting.

Example 7: Glucose uptake assay

For the determination of glucose uptake, cells were washed 3 times with PBS prior to serum starvation in Krebs-Ringer-Bicarbonat-Hepes buffer (KRBH; 134 nM NaCl, 3.5 mM KCl, 1.2 mM KH_2PO_4, 0.5 mM $MgSO_4$, 1.5 mM $CaCl_2$, 5 mM $NaHCO_3$, 10 mM Hepes, pH 7.4), supplemented with 0.1% FCS and 0.5mM Glucose for 2.5h at 37°C. For insulin-stimulated glucose uptake, cells were incubated with 1 microM bovine insulin (Sigma; carrier: 0.005N HCl) for 45 min at 37°C. Basal glucose uptake was determined with carrier only. Non-metabolizable 2-Deoxy-3H-D-Glucose (NEN Life Science, Boston, USA) in a final activity of 0,4 microCi/Well/ml was added for 30 min at 37°C. For the calculation of background radioactivity, 25 microM Cytochalasin B (Sigma) was used. All assays were performed in duplicate wells. To terminate the reaction, cells were washed 3 times with ice cold PBS, and lysed in 1 ml 0.1N NaOH. Protein

concentration of each well was assessed using the standard Biuret method (Protein assay reagent; Bio-Rad), and scintillation counting of cell lysates in 10 volumes Ultima-gold cocktail (Packard Bioscience, Groningen, Netherlands) was performed.

Example 8: Generation and analysis of NR1D1, NR1D2, Cdk4, Cdk6, CDC14A, and CDC14B transgenic mice

Generation of the transgenic animals

Mouse NR1D1, NR1D2, Cdk4, Cdk6, CDC14A, and CDC14B cDNA was isolated from mouse brown adipose tissue (BAT) using standard protocols as known to those skilled in the art. The cDNA was amplified by RT-PCR and point mutations were introduced into the cDNA.

The resulting mutated cDNA was cloned into a suitable transgenic expression vector. The transgene was microinjected into the male pronucleus of fertilized mouse embryos (preferably strain C57/BL6/CBA F1 (Harlan Winkelmann). Injected embryos were transferred into pseudo-pregnant foster mice. Transgenic founders were detected by PCR analysis. Two independent transgenic mouse lines containing the construct were established and kept on a C57/BL6 background. Briefly, founder animals were backcrossed with C57/BL6 mice to generate F1 mice for analysis. Transgenic mice were continously bred onto the C57/Bl6 background. The expression of the proteins of the invention can be analyzed by taqman analysis as described above, and further analysis of the mice can be done as known to those skilled in the art.

16. Dez. 2002

EPO - Munich
61
16. Dez. 2002

Claims

1. A pharmaceutical composition comprising a nucleic acid molecule encoding Eip75B, Cdk4, or CG7134 or Eip75B, Cdk4, or CG7134 homologs or a polypeptide encoded thereby or encoded by a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic acid molecule or said polypeptide, preferably together with pharmaceutically acceptable carriers and diluents.

2. The composition of claim 1, wherein the nucleic acid molecule is a vertebrate or insect Eip75B, Cdk4, or CG7134 nucleic acid, particulary encoding the human Eip75B, Cdk4, or CG7134 homologs (such as human NR1D1, NR1D2, CDK4, CDK6, CDC14A, CDC14B, MGC26484, or human protein similar to CDC14B isoform 3), and/or a nucleic molecule which is complementary thereto or a fragment thereof or a variant thereof.

3. The composition of claim 1 or 2, wherein said nucleic acid molecule is selected from the group consisting of
   (a) a nucleic acid molecule encoding a polypeptide as deposited under GenBank Accession Number NM_021724, NM_005126, NM_001259, NM_000075, NM_003672, NM_033312, NM_033313, NM_003671, NM_033331, NM_033332, NM_152627, or under Accession Number XM_171149, or an isoform, fragment or variant of the polypeptide as deposited under GenBank Accession Number NP_068370, NP_005117, NP_001250, NP_000066, NP_003663, NP_201569, NP_201570, NP_003662, NP_201588, NP_201589, NP_689840, or under Accession Number XP_171149;

(b) a nucleic acid molecule which comprises or is the nucleic acid molecule as deposited under GenBank Accession Number NM_021724, NM_005126, NM_001259, NM_000075, NM_003672, NM_033312, NM_033313, NM_003671, NM_033331, NM_033332, NM_152627, or under Accession Number XM_171149,

(c) a nucleic acid molecule being degenerate with as a result of the genetic code to a nucleic acid sequence as defined in (a) or (b),

(d) a nucleic acid molecule that hybridizes at 65°C in a solution containing 0.2 x SSC and 0.1% SDS to a nucleic acid molecule as defined in claim 2 or as defined in (a) to (c) and/or a nucleic acid molecule which is complementary thereto;

(e) a nucleic acid molecule that encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the human NR1D1, NR1D2, Cdk4, Cdk6, CDC14A, CDC14B, MGC26484, or similar to CDC14B isoform 3 variants, as defined in claim 2 or to a polypeptide as defined in (a);

(f) a nucleic acid molecule that differs from the nucleic acid molecule of (a) to (e) by mutation and wherein said mutation causes an alteration, deletion, duplication or premature stop in the encoded polypeptide.

4. The composition of any one of claims 1-3, wherein the nucleic acid molecule is a DNA molecule, particularly a cDNA or a genomic DNA.

5. The composition of any one of claims 1-4, wherein said nucleic acid encodes a polypeptide contributing to regulating the energy homeostasis and/or the metabolism of triglycerides.

6. The composition of any one of claims 1-5, wherein said nucleic acid molecule is a recombinant nucleic acid molecule.

7. The composition of any one of claims 1-6, wherein the nucleic acid molecule is a vector, particularly an expression vector.

8. The composition of any one of claims 1-5, wherein the polypeptide is a recombinant polypeptide.

9. The composition of claim 8, wherein said recombinant polypeptide is a fusion polypeptide.

10. The composition of any one of claims 1-7, wherein said nucleic acid molecule is selected from hybridization probes, primers and anti-sense oligonucleotides.

11. The composition of any one of claims 1-10 which is a diagnostic composition.

12. The composition of any one of claims 1-10 which is a therapeutic composition.

13. The composition of any one of claims 1-12 for the manufacture of an agent for detecting and/or verifying, for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and others, in cells, cell masses, organs and/or subjects.

14. Use of a nucleic acid molecule of the nuclear receptor subfamily 1 group D, cycline-dependent kinase, or cell division cycle 14 gene

family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic or polypeptide for controlling the function of a gene and/or a gene product which is influenced and/or modified by an Eip75B, Cdk4, or CG7134 homologous polypeptide.

15. Use of the nucleic acid molecule of the nuclear receptor subfamily 1 group D, cycline-dependent kinase, or cell division cycle 14 gene family or use of a nucleic acid molecule encoding Eip75B, Cdk4, or CG7134 or Eip75B, Cdk4, or CG7134 homologs or use ofa polypeptide encoded thereby, or use of a fragment or a variant of said nucleic acid molecule or said polypeptide, or use of an effector of said nucleic acid molecule or said polypeptide for identifying substances capable of interacting with an Eip75B, Cdk4, or CG7134 homologous polypeptide.

16. A non-human transgenic animal exhibiting a modified expression of an Eip75B, Cdk4, or CG7134 homologous polypeptide.

17. The animal of claim 16, wherein the expression of the Eip75B, Cdk4, or CG7134 homologous polypeptide is increased and/or reduced.

18. A recombinant host cell exhibiting a modified expression of an Eip75B, Cdk4, or CG7134 homologous polypeptide, or a recombinant host cell which comprises a nucleic acid molecule as defined in any one of claims 1 to 6.

19. The cell of claim 18 which is a human cell.

20. A method of identifying a (poly)peptide involved in the regulation of energy homeostasis and/or metabolism of triglycerides in a mammal comprising the steps of
    (a) contacting a collection of (poly)peptides with an Eip75B, Cdk4, or CG7134 homologous polypeptide or a fragment thereof under conditions that allow binding of said (poly)peptides;
    (b) removing (poly)peptides which do not bind and
    (c) identifying (poly)peptides that bind to said Eip75B, Cdk4, or CG7134 homologous polypeptide.

21. A method of screening for an agent which modulates the interaction of an Eip75B, Cdk4, or CG7134 homologous polypeptide with a binding target/agent, comprising the steps of
    (a) incubating a mixture comprising
        (aa) an Eip75B, Cdk4, or CG7134 homologous polypeptide or a fragment thereof;
        (ab) a binding target/agent of said Eip75B, Cdk4, or CG7134 homologous polypeptide or fragment thereof; and
        (ac) a candidate agent

        under conditions whereby said Eip75B, Cdk4, or CG7134 homologous polypeptide or fragment thereof specifically binds to said binding target/agent at a reference affinity;
    (b) detecting the binding affinity of said Eip75B, Cdk4, or CG7134 homologous polypeptide or fragment thereof to said binding target to determine an (candidate) agent-biased affinity; and
    (c) determining a difference between (candidate) agent-biased affinity and the reference affinity.

22. A method of screening for an agent, which modulates the activity of an Eip75B, Cdk4 or CG7134 homologous polypeptide comprising
    (a) incubating a mixture comprising
        (aa) an Eip75B, Cdk4, or CG7134 homologous polypeptide or a fragment thereof and
        (ab) a candidate agent

        under conditions whereby said Eip75B, Cdk4, or CG7134 homologous polypeptide or fragment thereof exhibits a reference activity,
    (b) detecting the activity of said Eip75B, Cdk4, or CG7134 homologous polypeptide or fragment thereof to determine an (candidate) agent-biased activity and
    (c) determining a difference between (candidate) agent-biased activity and reference activity.

23. A method of producing a composition comprising the (poly)peptide identified by the method of claim 20 or the agent identified by the method of claim 21 or 22 with a pharmaceutically acceptable carrier, diluent and/or adjuvant.

24. The method of claim 23 wherein said composition is a pharmaceutical composition for preventing, alleviating or treating of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, such as metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and other diseases and disorders.

25. Use of a (poly)peptide as identified by the method of claim 20 or of an agent as identified by the method of claim 21 or 22 for the preparation of a pharmaceutical composition for the treatment,

alleviation and/or prevention of of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and other diseases and disorders.

26. Use of a nucleic acid molecule as defined in any one of claims 1 to 6 or 10, use of a polypeptide as defined in any one of claims 1 to 6, 8 or 9, use of a vector as defined in claim 7, use of a host cell as defined in claim 18 or 19 for the preparation of a pharmaceutical composition for the treatment, alleviation and/or prevention of of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and other diseases and disorders.

27. Use of a nucleic acid molecule of the nuclear receptor subfamily 1 group D, cycline-dependent kinase, or cell division cycle 14 gene family or of a fragment thereof for the preparation of a non-human animal which over- or under-expresses the Eip75B, Cdk4, or CG7134 gene product.

28. Kit comprising at least one of
    (a) an Eip75B, Cdk4, or CG7134 nucleic acid molecule or a fragment or an isoform thereof;
    (b) an Eip75B, Cdk4, or CG7134 amino acid molecule or a fragment or an isoform thereof;
    (b) a vector comprising the nucleic acid of (a);
    (c) a host cell comprising the nucleic acid of (a) or the vector of (b);

(d) a polypeptide encoded by the nucleic acid of (a), expressed by the vector of (c) or the host cell of (a);

(e) a fusion polypeptide encoded by the nucleic acid of (a);

(f) an antibody, an aptamer or another receptor against the nucleic acid of (a) or the polypeptide of (b) , (e) , or (f) and /or

(g) an anti-sense oligonucleotide of the nucleic acid of (a).

Abstract

The present invention discloses Eip75B, Cdk4, or CG7134 homologous proteins regulating the energy homeostasis and the metabolism of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to the use of these sequences in the diagnosis, study, prevention, and treatment of metabolic diseases and disorders.

ld 16.12.2002

1124

FIGURE 1. Energy storage metabolite content of a Drosophila Eip75B mutant

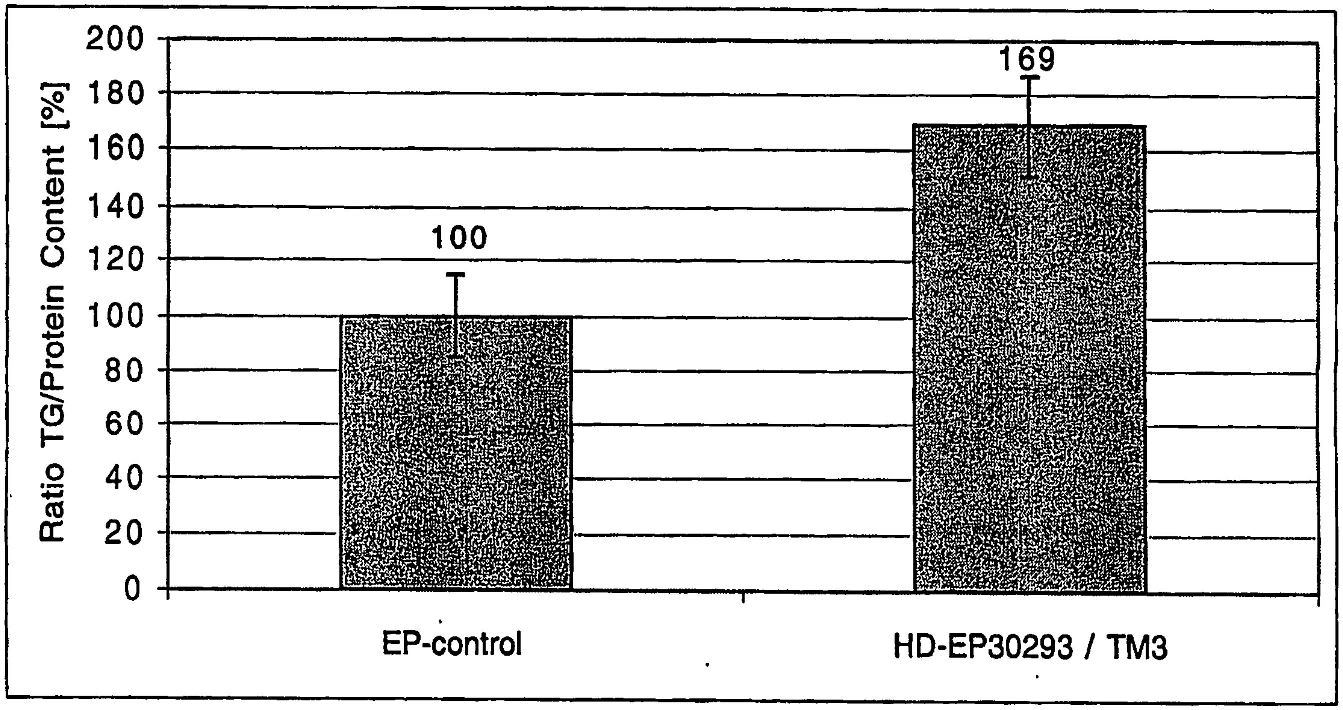


Figure 2. Molecular organization of the Eip75B gene (GadFly Acession Number CG8127)

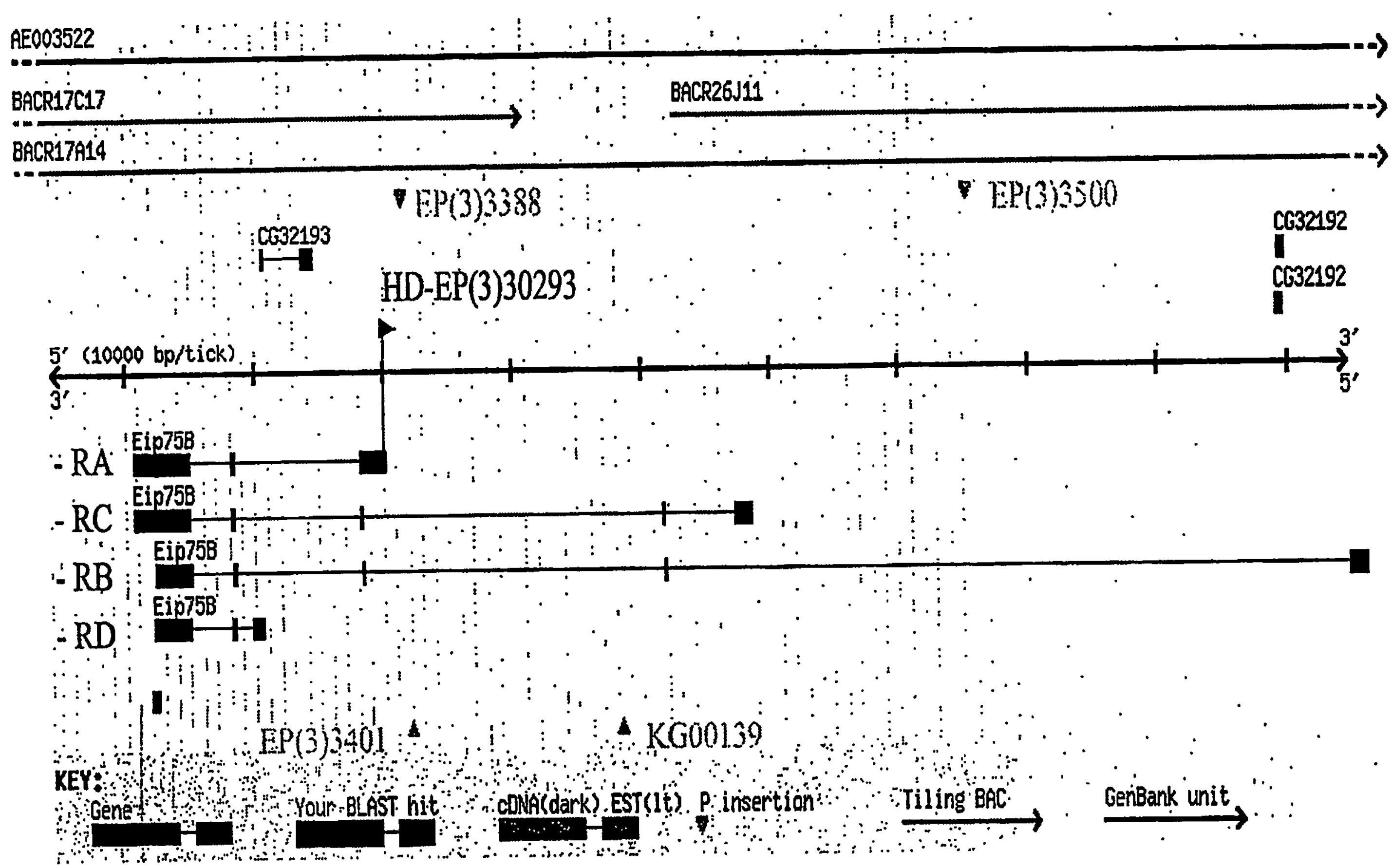


FIGURE 3. Nucleic acid sequences and amino acid sequences of the human nuclear receptor subfamily 1, group D

<u>FIGURE 3A.</u> Homo sapiens nuclear receptor subfamily 1, group D, member 1 (NR1D1), Nucleic acid sequence (SEQ ID NO: 1)

```
   1 ccgaggcgct ccctgggatc acatggtacc tgctccagtg ccgcgtgcgg cccgggaacc
  61 ctgggctgct ggcgcctgcg cagagccctc tgtcccaggg aaaggctcgg gcaaaaggcg
 121 gctgagattg gcagagtgaa atattactgc cgagggaacg tagcagggca cacgtctcgc
 181 ctctttgcga ctcggtgccc cgtttctccc catcacctac ttacttcctg gttgcaacct
 241 ctcttcctct gggactttg caccgggagc tccagattcg ctaccccgca gcgctgcgga
 301 gccggcaggc agaggcaccc cgtacactgc agagacccga ccctccttgc taccttctag
 361 ccagaactac tgcaggctga ttccccctac acactctctc tgctcttccc atgcaaagca
 421 gaactccgtt gcctcaacgt ccaacccttc tgcagggctg cagtccggcc accccaagac
 481 cttgctgcag ggtgcttcgg atcctgatcg tgagtcgcgg ggtccactcc ccgcccttag
 541 ccagtgccca gggggcaaca gcggcgatcg caacctctag tttgagtcaa ggtccagttt
 601 gaatgaccgc tctcagctgg tgaagacatg acgaccctgg actccaacaa caacacaggt
 661 ggcgtcatca cctacattgg ctccagtggc tcctccccaa gccgcaccag ccctgaatcc
 721 ctctatagtg acaactccaa tggcagcttc cagtccctga cccaaggctg tcccacctac
 781 ttcccaccat cccccactgg ctccctcacc caagacccgg ctcgctcctt tgggagcatt
 841 ccacccagcc tgagtgatga cggctcccct tcttcctcat cttcctcgtc gtcatcctcc
 901 tcctccttct ataatgggag ccccccctggg agtctacaag tggccatgga ggacagcagc
 961 cgagtgtccc ccagcaagag caccagcaac atcaccaagc tgaatggcat ggtgttactg
1021 tgtaaagtgt gtggggacgt tgcctcgggc ttccactacg gtgtgcacgc ctgcgagggc
1081 tgcaagggct ttttccgtcg gagcatccag cagaacatcc agtacaaaag gtgtctgaag
1141 aatgagaatt gctccatcgt ccgcatcaat cgcaaccgct gccagcaatg tcgcttcaag
1201 aagtgtctct ctgtgggcat gtctcgagac gctgtgcgtt ttgggcgcat ccccaaacga
1261 gagaagcagc ggatgcttgc tgagatgcag agtgccatga acctggccaa caaccagttg
1321 agcagccagt gcccgctgga gacttcaccc acccagcacc ccaccccagg ccccatgggc
1381 ccctcgccac cccctgctcc ggtcccctca cccctggtgg gcttctccca gtttccacaa
1441 cagctgacgc ctcccagatc ccaagccct gagcccacag tggaggatgt gatatcccag
1501 gtggcccggg cccatcgaga gatcttcacc tacgcccatg acaagctggg cagctcacct
1561 ggcaacttca atgccaacca tgcatcaggt agccctccag ccaccacccc acatcgctgg
1621 gaaaatcagg gctgcccacc tgcccccaat gacaacaaca ccttggctgc ccagcgtcat
1681 aacgaggccc taaatggtct gcgccaggct ccctcctcct accctcccac ctggcctcct
1741 ggccctgcac accacagctg ccaccagtcc aacagcaacg ggcaccgtct atgcccccacc
1801 cacgtgtatg cagccccaga aggcaaggca cctgccaaca gtccccggca gggcaactca
1861 aagaatgttc tgctggcatg tcctatgaac atgtacccgc atggacgcag tgggcgaacg
1921 gtgcaggaga tctgggagga tttctccatg agcttcacgc ccgctgtgcg ggaggtggta
1981 gagtttgcca aacacatccc gggcttccgt gacctttctc agcatgacca agtcaccctg
2041 cttaaggctg gcacctttga ggtgctgatg gtgcgctttg cttcgttgtt caacgtgaag
2101 gaccagacag tgatgttcct aagccgcacc acctacagcc tgcaggagct tggtgccatg
2161 ggcatgggag acctgctcag tgccatgttc gacttcagcg agaagctcaa ctccctggcg
2221 cttaccgagg aggagctggg cctcttcacc gcggtggtgc ttgtctctgc agaccgctcg
2281 ggcatggaga attccgcttc ggtggagcag ctccaggaga cgctgctgcg ggctcttcgg
2341 gctctggtgc tgaagaaccg gcccttggag acttcccgct tcaccaagct gctgctcaag
2401 ctgccggacc tgcggaccct gaacaacatg cattccgaga agctgctgtc cttccgggtg
2461 gacgcccagt gacccgcccg gccggccttc tgccgctgcc cccttgtaca gaatcgaact
2521 ctgcacttct ctctccttta cgagacgaaa aggaaaagca aaccagaatc ttatttatat
2581 tgttataaaa tattccaaga tgagcctctg gccccctgag ccttcttgta aatacctgcc
2641 tccctccccc atcaccgaac ttcccctcct cccctattta aaccactctg tctccccccac
2701 aaccctcccc tggccctctg atttgttctg ttcctgtctc aaatccaata gttcacagct
2761 gagctggg
```

FIGURE 3B. Homo sapiens nuclear receptor subfamily 1, group D, member 1 (NR1D1), Amino acid sequence (SEQ ID NO: 2)

```
  1 mttldsnnnt ggvityigss gsspsrtspe slysdnsngs fqsltqgcpt yfppsptgsl
 61 tqdparsfgs ippslsddgs psssssssss sssfyngspp gslqvameds srvspsksts
121 nitklngmvl lckvcgdvas gfhygvhace gckgffrrsi qqniqykrcl knencsivri
181 nrnrcqqcrf kkclsvgmsr davrfgripk rekqrmlaem qsamnlannq lssqcplets
241 ptqhptpgpm gpspppapvp splvgfsqfp qqltpprsps peptvedvis qvarahreif
301 tyahdklgss pgnfnanhas gsppattphr wenqgcppap ndnntlaaqr hnealnglrq
361 apssypptwp pgpahhschq snsnghrlcp thvyaapegk apansprqgn sknvllacpm
421 nmyphgrsgr tvqeiwedfs msftpavrev vefakhipgf rdlsqhdqvt llkagtfevl
481 mvrfaslfnv kdqtvmflsr ttyslqelga mgmgdllsam fdfseklnsl alteeelglf
541 tavvlvsadr sgmensasve qlqetllral ralvlknrpl etsrftklll klpdlrtlnn
601 mhsekllsfr vdaq
```

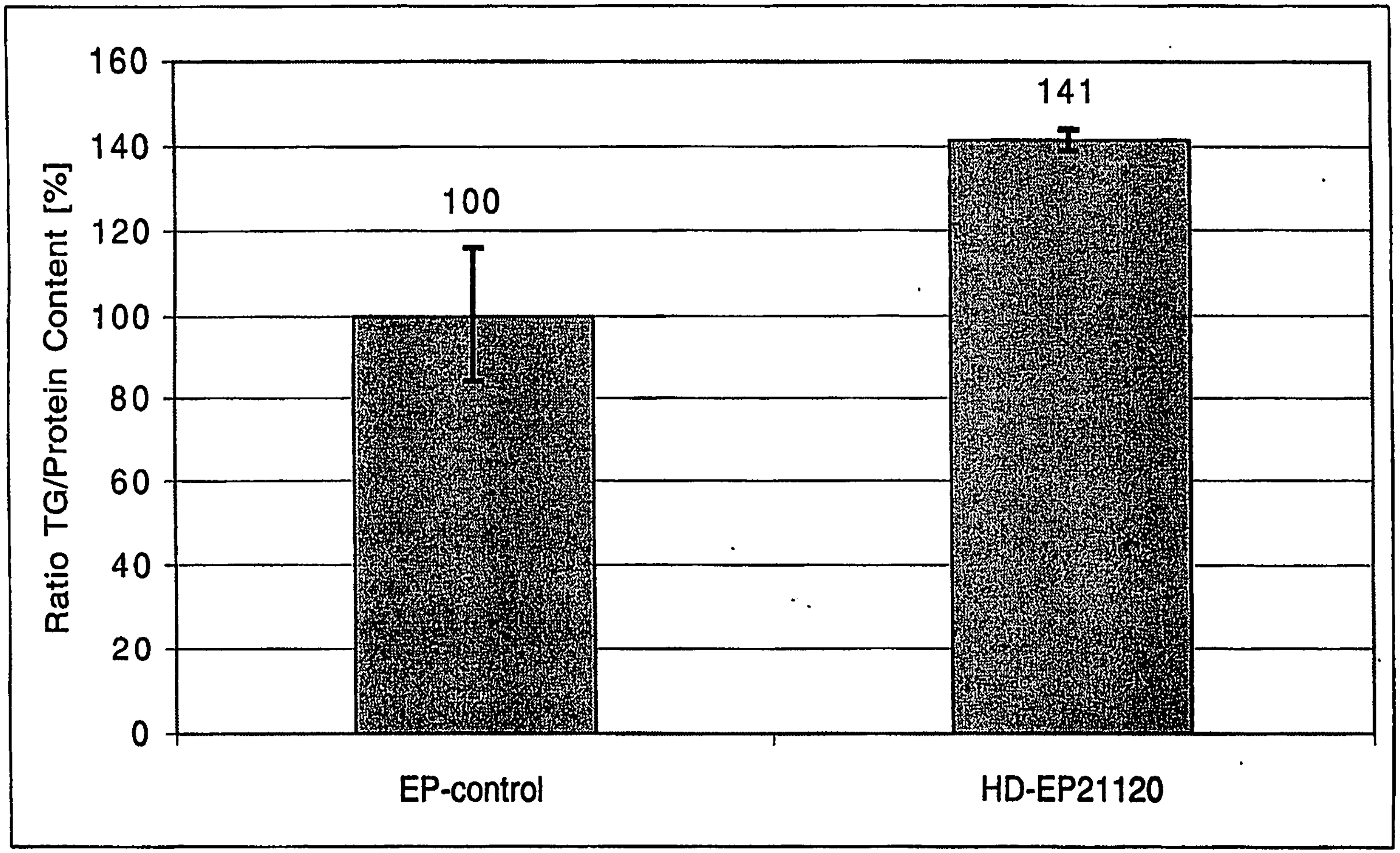
FIGURE 3C. Homo sapiens nuclear receptor subfamily 1, group D, member 2 (NR1D2), Nucleic acid sequence (SEQ ID NO: 3)

```
   1 gctgccctcc ccgtcagccg ccctcgccgc cgcggtgcgc tggctgcagg aagccgccgc
  61 gccgccgctt ttgttgtcag ggacccagcg aggagcgccg ctcgccggcc gccgccaccc
 121 tctctcgctg cagcctgctg tgcgctgcac ggcctggggc ccgggcgccc ccgcgtctgc
 181 ccatgagggg gccccgcgac caccgctgct tccagcccgg ggcggcgcgg cgctgaggcg
 241 gcggcggcgg cggcctgccc cctctgcggg aagcgggcgg ccccggccgc ctccgcgagg
 301 gcaccatgga ggtgaatgca ggaggtgtga ttgcctatat cagttcttcc agctcagcct
 361 caagccctgc ctcttgtcac agtgagggtt ctgagaatag tttccagtcc tcctcctctt
 421 ctgttccatc ttctccaaat agctctaatt ctgataccaa tggtaatccc aagaatggtg
 481 atctcgccaa tattgaaggc atcttgaaga atgatcgaat agattgttct atgaaaacaa
 541 gcaaatcgag tgcacctggg atgacaaaaa gtcatagtgg tgtgacaaaa tttagtggca
 601 tggttctact gtgtaaagtc tgtggggatg tggcgtcagg attccactat ggagttcatg
 661 cttgcgaagg ctgtaagggt ttctttcgga gaagtattca acaaaacatc cagtacaaga
 721 agtgcctgaa gaatgaaaac tgttctataa tgagaatgaa taggaacaga tgtcagcaat
 781 gtcgcttcaa aaagtgtctg tctgttggaa tgtcaagaga tgctgttcgg tttggtcgta
 841 ttcctaagcg tgaaaaacag aggatgctaa ttgaaatgca aagtgcaatg aagaccatga
 901 tgaacagcca gttcagtggt cacttgcaaa atgacacatt agtagaacat catgaacaga
 961 cagccttgcc agcccaggaa cagctgcgac ccaagcccca actggagcaa gaaaacatca
1021 aaagctcttc tcctccatct tctgattttg caaaggaaga agtgattggc atggtgacca
1081 gagctcacaa ggataccttt atgtataatc aagagcagca agaaaactca gctgagagca
1141 tgcagcccca gagaggagaa cggattccca agaacatgga gcaatataat ttaaatcatg
1201 atcattgcgg caatgggctt agcagccatt ttccctgtag tgagagccag cagcatctca
1261 atggacagtt caaagggagg aatataatgc attacccaaa tggtcatgcc atttgtattg
1321 caaatggaca ttgtatgaac ttctccaatg cttatactca aagagtatgt gatagagttc
1381 cgatagatgg atttctcag aatgagaaca agaatagtta cctgtgcaac actggaggaa
1441 gaatgcatct ggtttgtcca atgagtaagt ctccatatgt ggatcctcat aaatcaggac
1501 atgaaatctg ggaagaattt tcgatgagct tcactccagc agtgaaagaa gtggtggaat
1561 ttgcaaagcg tattcctggg ttcagagatc tctctcagca tgaccaggtc aacctttttaa
1621 aggctgggac ttttgaggtt ttaatggtac ggttcgcatc attatttgat gcaaaggaac
1681 gtactgtcac ctttttaagt ggaaagaaat atagtgtgga tgatttacac tcaatgggag
1741 caggggatct gctaaactct atgtttgaat ttagtgagaa gctaaatgcc ctccaactta
1801 gtgatgaaga gatgagtttg tttacagctg ttgtcctggt atctgcagat cgatctggaa
1861 tagaaaacgt caactctgtg gaggctttgc aggaaactct cattcgtgca ctaaggacct
1921 taataatgaa aaaccatcca aatgaggcct ctatttttac aaaactgctt ctaaagttgc
1981 cagatcttcg atctttaaac aacatgcact ctgaggagct cttggccttt aaagttcacc
2041 cttaaggcct ttgtttattt aaacatgaac tgatggtaac tgtacatttt gtgctaaaat
2101 gcatatttat atgtgcatac catatgtgga gatagaaaag accttta
```

FIGURE 3D. Homo sapiens nuclear receptor subfamily 1, group D, member 2 (NR1D2), Amino acid sequence (SEQ ID NO: 4)

```
  1 mevnaggvia yissssssass paschsegse nsfqsssssv psspnssnsd tngnpkngdl
 61 aniegilknd ridcsmktsk ssapgmtksh sgvtkfsgmv llckvcgdva sgfhygvhac
121 egckgffrrs iqqniqykkc lknencsimr mnrnrcqqcr fkkclsvgms rdavrfgrip
181 krekqrmlie mqsamktmmn sqfsghlqnd tlvehheqta lpaqeqlrpk pqleqeniks
241 ssppssdfak eevigmvtra hkdtfmynqe qqensaesmq pqrgeripkn meqynlnhdh
301 cgnglsshfp csesqqhlng qfkgrnimhy pnghaician ghcmnfsnay tqrvcdrvpi
361 dgfsqnenkn sylcntggrm hlvcpmsksp yvdphksghe iweefsmsft pavkevvefa
421 kripgfrdls qhdqvnllka gtfevlmvrf aslfdakert vtflsgkkys vddlhsmgag
481 dllnsmfefs eklnalqlsd eemslftavv lvsadrsgie nvnsvealqe tliralrtli
541 mknhpneasi ftklllklpd lrslnnmhse ellafkvhp
```

FIGURE 4. Energy storage metabolite content of a Drosophila Cdk4 mutant



6124

Figure 5. Molecular organization of the Cdk4 gene (GadFly Acession Number CG5072)

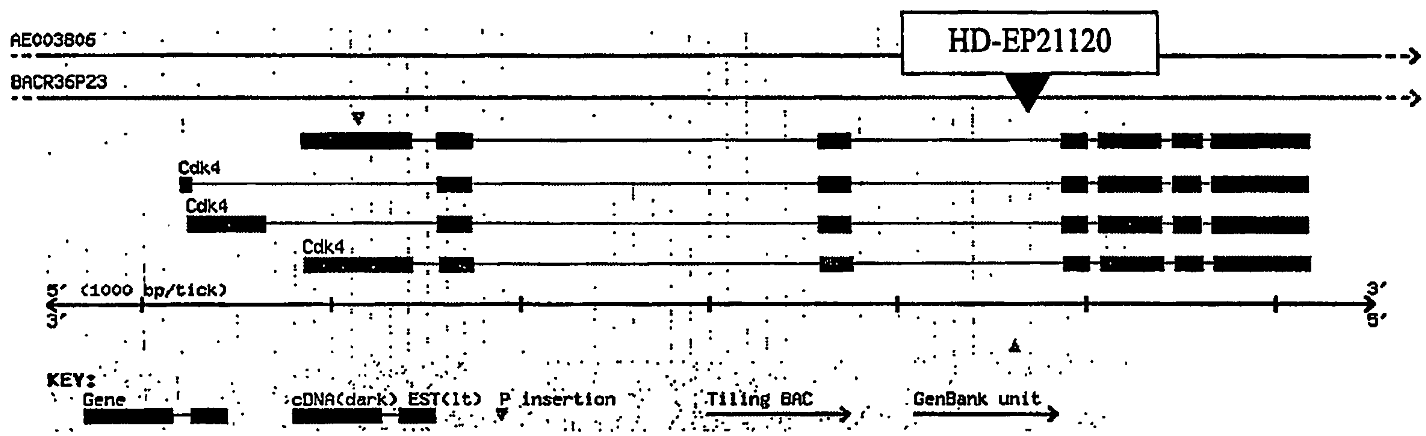


FIGURE 6. Nucleic acid sequences and amino acid sequences of the human cyclin-dependent kinases 4 and 6

FIGURE 6A. Homo sapiens cyclin-dependent kinase 6 (CDK6), Nucleic acid sequence (SEQ ID NO: 5)

```
   1 gtaaagctag accgatctcc ggggagcccc ggagtaggcg agcggcggcc gccagctagt
  61 tgagcgcacc ccccgcccgc cccagcggcg ccgcggcggg cggcgtccag gcggcatgga
 121 gaaggacggc ctgtgccgcg ctgaccagca gtacgaatgc gtggcggaga tcggggaggg
 181 cgcctatggg aaggtgttca aggcccgcga cttgaagaac ggaggccgtt tcgtggcgtt
 241 gaagcgcgtg cgggtgcaga ccggcgagga gggcatgccg ctctccacca tccgcgaggt
 301 ggcggtgctg aggcacctgg agaccttcga gcaccccaac gtggtcaggt tgtttgatgt
 361 gtgcacagtg tcacgaacag acagagaaac caaactaact ttagtgtttg aacatgtcga
 421 tcaagacttg accacttact tggataaagt tccagagcct ggagtgccca ctgaaaccat
 481 aaaggatatg atgtttcagc ttctccgagg tctggacttt cttcattcac accgagtagt
 541 gcatcgcgat ctaaaaccac agaacattct ggtgaccagc agcggacaaa taaaactcgc
 601 tgacttcggc cttgcccgca tctatagttt ccagatggct ctaacctcag tggtcgtcac
 661 gctgtggtac agagcacccg aagtcttgct ccagtccagc tacgccaccc ccgtggatct
 721 ctggagtgtt ggctgcatat ttgcagaaat gtttcgtaga aagcctcttt ttcgtggaag
 781 ttcagatgtt gatcaactag gaaaaatctt ggacgtgatt ggactcccag gagaagaaga
 841 ctggcctaga gatgttgccc ttcccaggca ggcttttcat tcaaaatctg cccaaccaat
 901 tgagaagttt gtaacagata tcgatgaact aggcaaagac ctacttctga agtgtttgac
 961 atttaaccca gccaaaagaa tatctgccta cagtgccctg tctcacccat acttccagga
1021 cctggaaagg tgcaaagaaa acctggattc ccacctgccg cccagccaga acacctcgga
1081 gctgaataca gcctgaggcc tcagcagccg ccttaagctg atcctgcgga gaacaccctt
1141 ggtggcttat gggtccccct cagcaagccc tacagagctg tggaggattg ctatctggag
1201 gccttccagc tgctgtcttc tggacaggct ctgcttctcc aaggaaa
```

FIGURE 6B. Homo sapiens cyclin-dependent kinase 6 (CDK6), Amino acid sequence (SEQ ID NO: 6)

```
  1 mekdglcrad qqyecvaeig egaygkvfka rdlknggrfv alkrvrvqtg eegmplstir
 61 evavlrhlet fehpnvvrlf dvctvsrtdr etkltlvfeh vdqdlttyld kvpepgvpte
121 tikdmmfqll rgldflhshr vvhrdlkpqn ilvtssgqik ladfglariy sfqmaltsvv
181 vtlwyrapev llqssyatpv dlwsvgcifa emfrrkplfr gssdvdqlgk ildviglpge
241 edwprdvalp rqafhsksaq piekfvtdid elgkdlllkc ltfnpakris aysalshpyf
301 qdlerckenl dshlppsqnt selnta
```

FIGURE 6C. Homo sapiens cyclin-dependent kinase 4 (CDK4), Nucleic acid sequence, isoform 1 (SEQ ID NO: 7)

```
  1 agccctccca gtttccgcgc gcctctttgg cagctggtca catggtgagg gtgggggtga
 61 gggggcctct ctagcttgcg gcctgtgtct atggtcgggc cctctgcgtc cagctgctcc
121 ggaccgagct cgggtgtatg ggccgtagg aaccggctcc ggggccccga taacgggccg
181 cccccacagc accccgggct ggcgtgaggg tctcccttga tctgagaatg gctacctctc
241 gatatgagcc agtggctgaa attggtgtcg gtgcctatgg gacagtgtac aaggcccgtg
301 atccccacag tggccacttt gtggccctca agagtgtgag agtccccaat ggaggaggag
361 gtggaggagg ccttcccatc agcacagttc gtgaggtggc tttactgagg cgactggagg
421 cttttgagca tcccaatgtt gtccggctga tggacgtctg tgccacatcc cgaactgacc
481 gggagatcaa ggtaaccctg gtgtttgagc atgtagacca ggacctaagg acatatctgg
541 acaaggcacc cccaccaggc ttgccagccg aaacgatcaa ggatctgatg cgccagtttc
601 taagaggcct agatttcctt catgccaatt gcatcgttca ccgagatctg aagccagaga
661 acattctggt gacaagtggt ggaacagtca agctggctga ctttggcctg gccagaatct
```

```
 721 acagctacca gatggcactt acacccgtgg ttgttacact ctggtaccga gctcccgaag
 781 ttcttctgca gtccacatat gcaacacctg tggacatgtg gagtgttggc tgtatctttg
 841 cagagatgtt tcgtcgaaag cctctcttct gtggaaactc tgaagccgac cagttgggca
 901 aaatctttga cctgattggg ctgcctccag aggatgactg gcctcgagat gtatccctgc
 961 cccgtggagc ctttcccccc agagggcccc gcccagtgca gtcggtggta cctgagatgg
1021 aggagtcggg agcacagctg ctgctggaaa tgctgacttt taacccacac aagcgaatct
1081 ctgcctttcg agctctgcag cactcttatc tacataagga tgaaggtaat ccggagtgag
1141 caatggagtg gctgccatgg aaggaagaaa agctgccatt tcccttctgg acactgagag
1201 ggcaatcttt gcctttatct ctgaggctat ggagggtcct cctccatctt tctacagaga
1261 ttactttgct gccttaatga cattcccctc ccacctctcc ttttgaggct tctccttctc
1321 cttcccattt ctctacacta aggggtatgt tccctcttgt ccctttccct acctttatat
1381 ttggggtcct tttttataca ggaaaaacaa aacaaagaaa taatggtctt tttttttttt
1441 ttaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa
```

<u>FIGURE 6D</u>. Homo sapiens cyclin-dependent kinase 4 (CDK4), Amino acid sequence, isoform 1 (SEQ ID NO: 8)

```
  1 matsryepva eigvgaygtv ykardphsgh fvalksvrvp nggggggglp istvrevall
 61 rrleafehpn vvrlmdvcat srtdreikvt lvfehvdqdl rtyldkappp glpaetikdl
121 mrqflrgldf lhancivhrd lkpenilvts ggtvkladfg lariysyqma ltpvvvtlwy
181 rapevllqst yatpvdmwsv gcifaemfrr kplfcgnsea dqlgkifdli glppeddwpr
241 dvslprgafp prgprpvqsv vpemeesgaq lllemltfnp hkrisafral qhsylhkdeg
301 npe
```

FIGURE 7. Energy storage metabolite content of a Drosophila CC7134 (Gadfly Accession Number) mutant

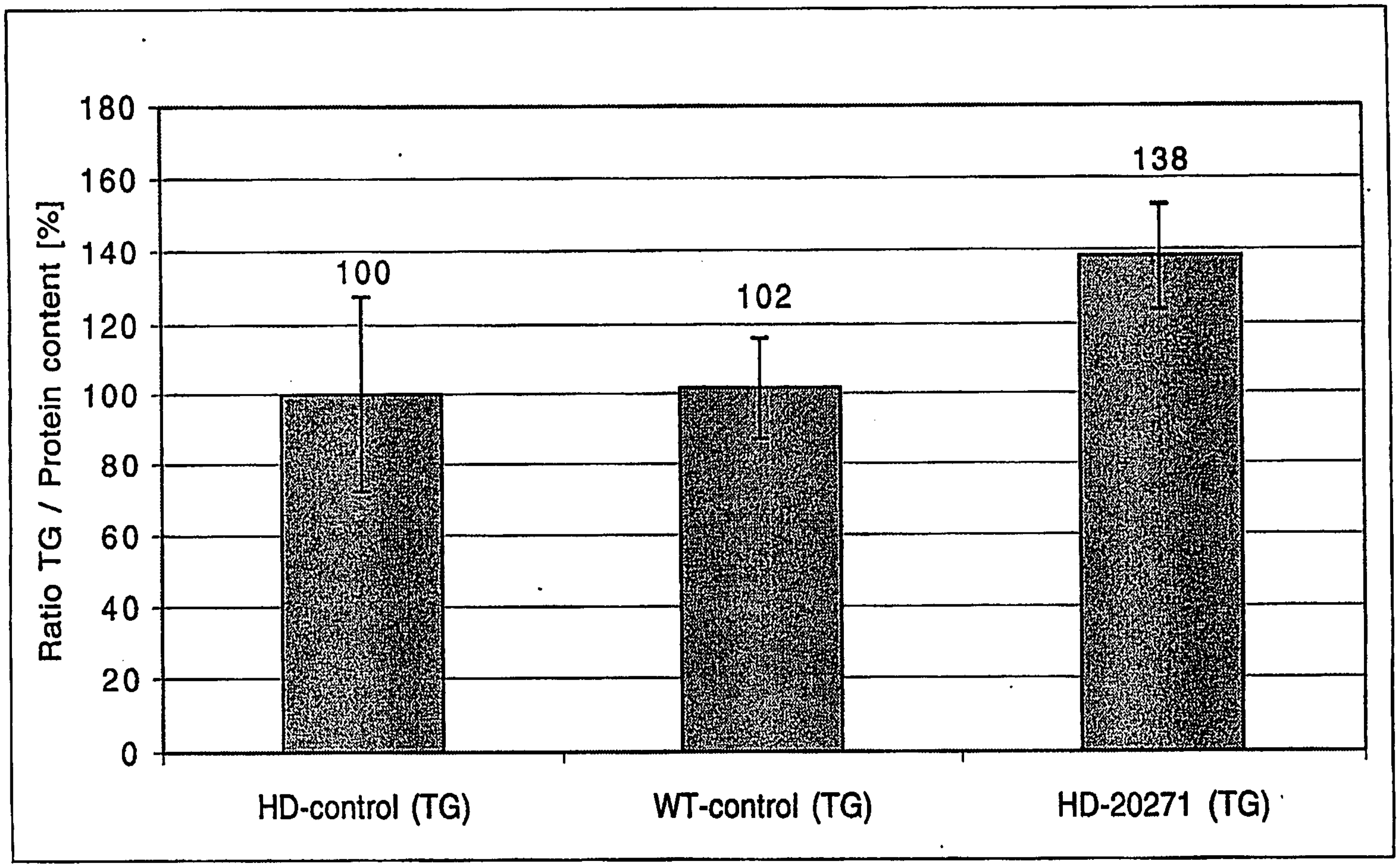


Figure 8. Molecular organization of the CG7134 gene (GadFly Acession Number)

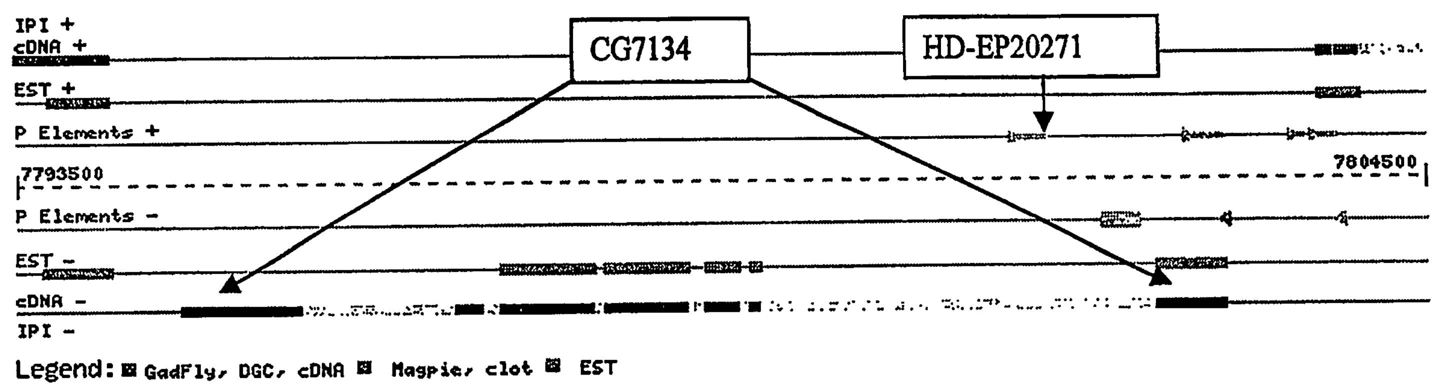


FIGURE 9. Nucleic acid sequences and amino acid sequences of the human cell division cycle 14 homologs A, B, and MGC26484

FIGURE 9A. Homo sapiens CDC14 homolog A, Nucleic acid sequence, transcript variant 1 (SEQ ID NO: 9)

```
   1 cgaagaggat ccggagcagc tgctgccagc ccgcgggcac tgaagtcctc ccggctgccg
  61 ctcgagtagc cacgggcgcg atcgggacca gaagtctcct cctccatgat cactttggaa
 121 gccggggggg aagactttgc cctgccctga gagctggtct gcgtttccca ggcgcggcgg
 181 cggcggagca gcagctgcag cagccgagtc caaataggag cggccacagc caggggcgtg
 241 tgcgccccgc gcggagcgag ctcgggttcc cctcggaatg tccccggggc gcccggcgcg
 301 ctgaccccga agccgcctcc gccttcggcg cctgctgcct ccctcggcca ggcttgttgt
 361 tcgggactgt gagcttcctg gctcctgggc agtggggaag cccccggggg cgagtgacct
 421 cagctggcca cgacccagcc ctcccccgtg cgtatctcgc ttaagatggc agcggagtca
 481 ggggaactaa tcggggcttg tgagttcatg aaagatcggt tatattttgc tactttaagg
 541 aatagaccaa aaagcacagt aaatacccac tatttctcca tcgatgagga gctggtctat
 601 gaaaatttct atgcagattt tggaccgctg aacttggcaa tggtgtacag atattgctgc
 661 aaactaaaca agaaactaaa atcatacagt ttgtcaagaa agaaaatagt gcactacacc
 721 tgttttgacc aacggaaaag agcaaatgca gcatttttga taggtgccta tgcagtaatc
 781 tatttaaaga agacaccaga agaagcctac agagcactcc tgtctggctc aaaccccccc
 841 tatcttccat tcagggatgc ttcctttgga aattgcactt acaatctcac cattctcgac
 901 tgtttgcagg gaatcagaaa gggattacaa catggatttt ttgactttga gacatttgat
 961 gtggatgaat atgaacatta tgagcgagtt gaaaatggtg acttcaactg gattgttcca
1021 ggaaaatttt tagcatttag tggaccacat cctaaaagca aattgagaa tggttatcct
1081 cttcacgccc ctgaagccta ctttccttat ttcaaaaagc ataatgtgac tgcagttgtg
1141 aggctaaaca aaaagattta tgaggcaaag cgcttcacag acgctggctt cgagcactat
1201 gacctcttct tcatagatgg cagcacaccc agtgacaaca tcgtgcgaag gttcctgaac
1261 atctgtgaga acaccgaagg ggccatcgcc gttcactgca aagctggtct tggaagaaca
1321 gggacattga tagcctgtta tgtaatgaaa cactacaggt ttacacatgc tgaaataatt
1381 gcttggatta gaatatgccg gccaggctct attataggac cccagcagca cttcctggaa
1441 gaaaaacaag catcgttgtg ggtccaagga gacattttcc gatccaaact gaaaaatcga
1501 ccatccagtg aaggaagtat taataaaatt ctttctggcc tagatgatat gtctattggt
1561 ggaaatcttt caaaaacaca aaacatggaa cgatttggag aggataactt agaagatgat
1621 gatgtggaaa tgaaaaatgg tataacccag ggagacaaac tacgtgcctt aaaaagtcag
1681 agacagccac gtacctcacc atcctgtgca tttaggtcag atgatacaaa aggacatcca
1741 agagcagtgt cccagccttt cagattaagt tcatccctgc aaggatctgc agttactttg
1801 aagacatcaa aaatggcact gtccccttca gcaacggcca agaggatcaa cagaacttct
1861 ttgtcttcgg gtgccactgt aagaagcttt tccataaact cccggctagc cagttctcta
1921 gggaacttga atgctgcaac agatgatcca gagaacaaaa agacctcctc atcctctaag
1981 gcaggcttca cagccagccc gtttaccaac ctcttgaatg gcagctccca gccaactacc
2041 agaaattacc ctgagctcaa caataatcag tacaacagaa gcagcaacag caacgggggc
2101 aacctgaaca gccccccagg cccccacagc gccaagacag aggagcacac caccatcctc
2161 cgaccctcct acaccgggct ttcttcttct tcagcgagat tcctgagccg ttctatccct
2221 tcccttcagt ctgaatatgt tcattactaa ggccttgcca ctccagtgaa agctgttctt
2281 ctcttagaca caatttcttc atctggacga gcagtggaga gggaaagcaa cttcttgctg
2341 gaagaatatc tctgccttct taccttaaat taaaaagagc actaagataa caccttcaag
2401 agacttgaaa acagaaaact ggttaatgac tactataaat gcactgaaac tatgttatgg
2461 agatttccat acttttaaag acagttttaa tgttgaattt ggtattttga agggttattt
2521 ttaatgtatt ttggtaatac atttattatt atatttacat gtacagtgtt acattatata
2581 tgtattgtga actttaaaag actattttga taaatttata aatatataaa attatgtaaa
2641 aactacacta tattttgatt tagattttcc tgctgtttgc taccaaaaat ttgtatttta
2701 aatctgttta gttttagtat ggttttgtct ctaatgaata aataattcct tcttattaag
2761 aagaagtaag ggagaaagtt tttagaaagt gatttttatg ctcgcactat aaatatggca
2821 ggtcagttca ttcttttggg aagtcagttt agttacactg agtttatcca agtttatctc
2881 taccaagagt ataatggcat gggatggctt atttaggaca attccctttc ctattgtttt
2941 tgttgctgag ccaatttgag ttagttttgc atcctggggg gctttaaaat acagcatgca
3001 gtgaaagatc agaattcact gaatatttct tctgagagca tggtttcatg gtttttctct
```

3061 atgaaatgac tcaatattcc aaatgttttt ttttccttcc tcctttcaaa agagttctta
3121 acccaattag gatatcctgc tttgggtatg aggttgttgt tgcctgtaat cacacatggt
3181 ttgacatcag ttttaaatca atggagagaa aaaactgaaa aagatgctgc taagtagttc
3241 tctgtattaa aggagatatt tttaaaacag ggtacaaccc cctgctgcac acgctagcat
3301 atctggaacc tactatgaaa atgaaaggac ccttataggt actcacagcc ctttcatgta
3361 agtatgatct gatatttagg tcttcagaag cctgtaggtt tcatttctat gaggaatcga
3421 ggagcgttac atcctgatat ccttccaggc tgcttaagaa tggactgctt cgacactgaa
3481 agtgctagtt aaatggattc atatgaagtg ctttactccc aaccattgag ttatttataa
3541 tgtatttatt aggggagggt accttgagtc tattatatat gcttcatcaa aacatcttgt
3601 tcatgtttta tgtttttaaa aaaggcattt gaatgaatgt ttgactcagg tttgttaaat
3661 taacttcagt aactgcagta ccaaaaatta cactcaactg atgaaaaaaa cgaattgtat
3721 gatttaggaa tcaaaaacta aaataagtgg aattatgtat cttttctaaa gttaaaaaag
3781 taaaatattt tattatgagt tattataaaa attggttaat tgtataggaa gatgacagta
3841 tttttttcaa gttatcataa aaagtaattc agatgacatt tgagaagtag gggaaaggga
3901 atcatgttga cagttttagt tctgtgaaca ctaatttgtg tgaagctatt aaaatgattg
3961 taaagttgac tactgtaaat ttcccataat tatgtgtgta tatgtgtcat atgtatgtac
4021 atgtatatgt ctaaaaatta ctttacacat gtgcctacat agacacacca agaagtggat
4081 gtatataata tagaaagtat atagcaaagt aattttactc tgataataaa aattgtttga
4141 catgtatttt gttatgaata gtttatcttc caaaagatat tttgctctat tttaaagtgt
4201 agaagaatac actgctaata aataataaaa gttttattca atttaaaaaa aaaaaaaaaa
4261 aa

FIGURE 9B. Homo sapiens CDC14 homolog A, Amino acid sequence, isoform 1 (SEQ ID NO: 10)

1 maaesgelig acefmkdrly fatlrnrpks tvnthyfsid eelvyenfya dfgplnlamv
61 yryccklnkk lksyslsrkk ivhytcfdqr kranaaflig ayaviylkkt peeayralls
121 gsnppylpfr dasfgnctyn ltildclqgi rkglqhgffd fetfdvdeye hyervengdf
181 nwivpgkfla fsgphpkski engyplhape ayfpyfkkhn vtavvrlnkk iyeakrftda
241 gfehydlffi dgstpsdniv rrflnicent egaiavhcka glgrtgtlia cyvmkhyrft
301 haeiiawiri crpgsiigpq qhfleekqas lwvqgdifrs klknrpsseg sinkilsgld
361 dmsiggnlsk tqnmerfged nledddvemk ngitqgdklr alksqrqprt spscafrsdd
421 tkghpravsq pfrlssslqg savtlktskm alspsatakr inrtslssga tvrsfsinsr
481 lasslgnlna atddpenkkt sssskagfta spftnllngs sqpttrnype lnnnqynrss
541 nsnggnlnsp pgphsaktee httilrpsyt glssssarfl srsipslqse yvhy

FIGURE 9C. Homo sapiens CDC14 homolog A, Nucleic acid sequence, transcript variant 2 (SEQ ID NO: 11)

1 cgaagaggat ccggagcagc tgctgccagc ccgcgggcac tgaagtcctc ccggctgccg
61 ctcgagtagc cacgggcgcg atcgggacca gaagtctcct cctccatgat cactttggaa
121 gccggggggaa gactttgccc tgccctgaga gctggtctgc gtttcccagg cgcggcggcg
181 gcggagcagc agctgcagca gccgagtcca aataggagcg gccacagcca ggggcgtgtg
241 cgccccgcgc ggagcgagct cgggttcccc tcggaatgtc cccggggcgc ccggcgcgct
301 gaccccgaag ccgcctccgc cttcggcgcc tgctgcctcc ctcggccagg cttgttgttc
361 gggactgtga gcttcctggc tcctgggcag tggggaagcc cccgggggcg agtgacctca
421 gctggccacg acccagccct cccccgtgcg tatctcgctt aagatggcag cggagtcagg
481 ggaactaatc ggggcttgtg agttcatgaa agatcggtta tattttgcta ctttaaggaa
541 tagaccaaaa agcacagtaa atacccacta tttctccatc gatgaggagc tggtctatga
601 aaatttctat gcagattttg gaccgctgaa cttggcaatg gtgtacagat attgctgcaa
661 actaaacaag aaactaaaat catacagttt gtcaagaaag aaaatagtgc actacacctg
721 ttttgaccaa cggaaaagag caaatgcagc atttttgata ggtgcctatg cagtaatcta
781 tttaaagaag acaccagaag aagcctacag agcactcctg tctggctcaa acccccccta
841 tcttccattc agggatgctt cctttggaaa ttgcacttac aatctcacca ttctcgactg

901 tttgcaggga atcagaaagg gattacaaca tggatttttt gactttgaga catttgatgt
961 ggatgaatat gaacattatg agcgagttga aaatggtgac ttcaactgga ttgttccagg
1021 aaaattttta gcatttagtg gaccacatcc taaaagcaaa attgagaatg gttatcctct
1081 tcacgcccct gaagcctact ttccttattt caaaaagcat aatgtgactg cagttgtgag
1141 gctaaacaaa aagatttatg aggcaaagcg cttcacagac gctggcttcg agcactatga
1201 cctcttcttc atagatggca gcacacccag tgacaacatc gtgcgaaggt tcctgaacat
1261 ctgtgagaac accgaagggg ccatcgccgt tcactgcaaa gctggtcttg gaagaacagg
1321 gacattgata gcctgttatg taatgaaaca ctacaggttt acacatgctg aaataattgc
1381 ttggattaga atatgccggc caggctctat tataggaccc cagcagcact tcctggaaga
1441 aaaacaagca tcgttgtggg tccaaggaga cattttccga tccaaactga aaaatcgacc
1501 atccagtgaa ggaagtatta ataaaattct ttctggccta gatgatatgt ctattggtgg
1561 aaatctttca aaaacacaaa acatggaacg atttggagag gataacttag aagatgatga
1621 tgtggaaatg aaaaatggta taacccaggg agacaaacta cgtgccttaa aaagtcagag
1681 acagccacgt acctcaccat cctgtgcatt taggtcagat gatacaaaag gacatccaag
1741 agcagtgtcc cagcctttca gattaagttc atccctgcaa ggatctgcag ttactttgaa
1801 gacatcaaaa atggcactgt ccccttcagc aacggccaag aggatcaaca gaacttcttt
1861 gtcttcgggt gccactgtaa gaagcttttc cataaactcc cggctagcca gttctctagg
1921 gaacttgaat gctgcaacag atgatccaga gaacaaaaag acctcctcat cctctaaggc
1981 aggcttcaca gccagcccgt ttaccaacct cttgaatggc agctcccagc caactaccag
2041 aaattaccct gagctcaaca ataatcagta caacagaagc agcaacagca acgggggcaa
2101 cctgaacagc cccccaggcc cccacagcgc caagacagag gagcacacca ccatcctccg
2161 accctcctac accgggcttt cttcttcttc agcgagattc ctgagccgtt ctatccctgt
2221 aagtgcgcag acaccacctc ctggtcctca gaaccctgaa tgcaacttct gtgccttgcc
2281 ttcccagccg aggctgccac caaagaaatt taatagtgcc aaggaagcct tctgagcgat
2341 gccttccctc tgtgctgtga aactgtctat gcactacatt ctgctagctc ctcttcaagt
2401 aaacgccaag tcacaaaaaa aaaaaaaaaa aaaaaaaa

<u>FIGURE 9D</u>. Homo sapiens CDC14 homolog A, Amino acid sequence, isoform 2 (SEQ ID NO: 12)

1 maaesgelig acefmkdrly fatlrnrpks tvnthyfsid eelvyenfya dfgplnlamv
61 yrycckl nkk lksyslsrkk ivhytcfdqr kranaaflig ayaviylkkt peeayralls
121 gsnppylpfr dasfgnctyn ltildclqgi rkglqhgffd fetfdvdeye hyervengdf
181 nwivpgkfla fsgphpkski engyplhape ayfpyfkkhn vtavvrlnkk iyeakrftda
241 gfehydlffi dgstpsdniv rrflnicent egaiavhcka glgrtgtlia cyvmkhyrft
301 haeiiawiri crpgsiigpq qhfleekqas lwvqgdifrs klknrpsseg sinkilsgld
361 dmsiggnlsk tqnmerfged nledddvemk ngitqgdklr alksqrqprt spscafrsdd
421 tkghpravsq pfrlssslqg savtlktskm alspsatakr inrtslssga tvrsfsinsr
481 lasslgnlna atddpenkkt ssssk agfta spftnllngs sqpttrnype lnnnqynrss
541 nsnggnlnsp pgphsaktee httilrpsyt glssssarfl srsipvsaqt pppgpqnpec
601 nfcalpsqpr lppkkfnsak eaf

<u>FIGURE 9E</u>. Homo sapiens CDC14 homolog A, Nucleic acid sequence, transcript variant 3 (SEQ ID NO: 13)

1 cgaagaggat ccggagcagc tgctgccagc ccgcgggcac tgaagtcctc ccggctgccg
61 ctcgagtagc cacgggcgcg atcgggacca gaagtctcct cctccatgat cactttggaa
121 gccgggggaa gactttgccc tgccctgaga gctggtctgc gtttcccagg cgcggcggcg
181 gcggagcagc agctgcagca gccgagtcca aataggagcg gccacagcca ggggcgtgtg
241 cgccccgcgc ggagcgagct cgggttcccc tcggaatgtc cccggggcgc ccggcgcgct
301 gaccccgaag ccgcctccgc cttcggcgcc tgctgcctcc ctcggccagg cttgttgttc
361 gggactgtga gcttcctggc tcctgggcag tggggaagcc cccgggggcg agtgacctca
421 gctggccacg acccagccct cccccgtgcg tatctcgctt aagatggcag cggagtcagg
481 ggaactaatc ggggcttgtg agttcatgaa agatcggtta tattttgcta ctttaaggaa

541 tagaccaaaa agcacagtaa atacccacta tttctccatc gatgaggagc tggtctatga
601 aaatttctat gcagattttg gaccgctgaa cttggcaatg gtgtacagat attgctgcaa
661 actaaacaag aaactaaaat catacagttt gtcaagaaag aaaatagtgc actacacctg
721 ttttgaccaa cggaaaagag caaatgcagc atttttgata ggtgcctatg cagtaatcta
781 tttaaagaag acaccagaag aagcctacag agcactcctg tctggctcaa acccccccta
841 tcttccattc agggatgctt cctttggaaa ttgcacttac aatctcacca ttctcgactg
901 tttgcaggga atcagaaagg gattacaaca tggatttttt gactttgaga catttgatgt
961 ggatgaatat gaacattatg agcgagttga aaatggtgac ttcaactgga ttgttccagg
1021 aaaattttta gcatttagtg gaccacatcc taaaagcaaa attgagaatg gttatcctct
1081 tcacgcccct gaagcctact ttccttattt caaaaagcat aatgtgactg cagttgtgag
1141 gctaaacaaa aagatttatg aggcaaagcg cttcacagac gctggcttcg agcactatga
1201 cctcttcttc atagatggca gcacacccag tgacaacatc gtgcgaaggt tcctgaacat
1261 ctgtgagaac accgaagggg ccatcgccgt tcactgcaaa gctggtcttg gaagaacagg
1321 gacattgata gcctgttatg taatgaaaca ctacaggttt acacatgctg aaataattgc
1381 ttggattaga atatgccggc caggctctat tataggaccc cagcagcact tcctggaaga
1441 aaaacaagca tcgttgtggg tccaaggaga cattttccga tccaaactga aaaatcgacc
1501 atccagtgaa ggaagtatta ataaaattct ttctggccta gatgatatgt ctattggtgg
1561 aaatctttca aaaacacaaa acatggaacg atttggagag gtaagttttc cctaggagat
1621 tctatcttct taaaactgat gttctgcatt tgtttctcag ttggacctat ataacatagc
1681 agtgtctttt ctctggatgc cacgagtacc aagtttttag aagtagagcc atccgtctat
1741 atagcaagaa gcagaggaaa gaaaccaatt gcccttaaaa aaaaaaagct ataatttaag
1801 gagtaaatta taaaggaggc tactctggta aggggtaata tttatagaaa ggaaacagaa
1861 aagcaaactt tctatttgaa aaaaaaaaaa aaaaaaaaaa

FIGURE 9F. Homo sapiens CDC14 homolog A, Amino acid sequence, isoform 3 (SEQ ID NO: 14)

1 maaesgelig acefmkdrly fatlrnrpks tvnthyfsid eelvyenfya dfgplnlamv
61 yryccklnkk lksyslsrkk ivhytcfdqr kranaaflig ayaviylkkt peeayralls
121 gsnppylpfr dasfgnctyn ltildclqgi rkglqhgffd fetfdvdeye hyervengdf
181 nwivpgkfla fsgphpkski engyplhape ayfpyfkkhn vtavvrlnkk iyeakrftda
241 gfehydlffi dgstpsdniv rrflnicent egaiavhcka glgrtgtlia cyvmkhyrft
301 haeiiawiri crpgsiigpq qhfleekqas lwvqgdifrs klknrpsseg sinkilsgld
361 dmsiggnlsk tqnmerfgev sfp

FIGURE 9G. Homo sapiens CDC14 homolog B, Nucleic acid sequence, transcript variant 1 (SEQ ID NO: 15)

1 cacggaacag ccctcctggg gtccccacga gccgcgtcct gctgtgcccc ggcgcctacg
61 cagcagcggc cgcggccgcg gtgggcacgc acggttaccc cgggcagctc cggccgccag
121 ctgcagcccc gtcgcctcgg ccgcgccagc cggctgcggg cacctggggg cgggctgggg
181 gcgccggccg cggcaggagg cgctgtagcg agggctgcgg cgccggtcct gcggcggccg
241 cgggaggcag cggggcaggc gctgtgggcc gggctcctcc tccggctcct gcgcgaccgc
301 ctcccgccgg gctctgccgg cgcccgccgt ccccgcagcg ccgctctgcg cccgccgccc
361 cgagcgcccg cgcggggctg gcgggagcct cggcgggcgc gcgggcgcgc ggggccatgg
421 tcgtggcccc ctgacgggcc gcggccgcct ccatgaagcg gaaaagcgag cggcggtcga
481 gctgggccgc cgcgcccccc tgctcgcggc gctgctcgtc gacctcgccg ggtgtgaaga
541 agatccgcag ctccacgcag caagacccgc gccgccggga cccccaggac gacgtgtacc
601 tggacatcac cgatcgcctt tgttttgcca ttctctacag cagaccaaag agtgcatcaa
661 atgtacatta tttcagcata gataatgaac ttgaatatga gaacttctac gcagattttg
721 gaccactcaa tctggcaatg gtttacagat attgttgcaa gatcaataag aaattaaagt
781 ccattacaat gttaaggaag aaaattgttc attttactgg ctctgatcag agaaaacaag
841 caaatgctgc cttccttgtt ggatgctaca tggttatata tttggggaga accccagaag
901 aagcatatag aatattaatc tttggagaga catcctatat tcctttcaga gatgctgcct

961 atggaagttg caatttctac attacacttc ttgactgttt tcatgcagta aagaaggcaa
1021 tgcagtatgg cttccttaat ttcaactcat ttaaccttga tgaatatgaa cactatgaaa
1081 aagcagaaaa tggagattta aattggataa taccagaccg atttattgcc ttctgtggac
1141 ctcattcaag agccagactt gaaagtggtt accaccaaca ttctcctgag acttatattc
1201 aatattttaa gaatcacaat gttactacca ttattcgtct gaataaaagg atgtatgatg
1261 ccaaacgctt tacggatgct ggcttcgatc accatgatct ttctttgcg gatggcagca
1321 cccctactga tgccattgtc aaagaattcc tagatatctg tgaaaatgct gagggtgcca
1381 ttgcagtaca ttgcaaagct ggccttggtc gcacgggcac tctgatagcc tgctacatca
1441 tgaagcatta caggatgaca gcagccgaga ccattgcgtg ggtcaggatc tgcagacctg
1501 gctcggtgat tgggcctcag cagcagtttt tggtgatgaa gcaaaccaac ctctggctgg
1561 aaggggacta ttttcgtcag aagttaaagg ggcaggagaa tggacaacac agagcagcct
1621 tctccaaact tctctctggc gttgatgaca tttccataaa tggggtcgag aatcaagatc
1681 agcaagaacc cgaaccgtac agtgatgatg acgaaatcaa tggagtgaca caaggtgata
1741 gacttcgggc cttgaaaagc agaagacaat ccaaaacaaa cgctattcct ctcactctct
1801 ccatttcaag gactaaaaca gtcttgcgtt aagtaaaaac ctgtgaccag agctgaagga
1861 agactctagg actgaaaact gcaacagaaa ttagcacaat ttgaaaacaa aacaaaattg
1921 caaaagcctt agttgctttt tccacctaag aagttgatca atggagaaaa tgtccactgg
1981 agtttgaata atgaactttg agtttgggtg caagcaaatg actcagagaa gggtccagct
2041 ctcaagctga atgacaaaca tgctgttgta aatttagtct caggtgtaaa tacccaagcc
2101 ctctggtacc cagggagctg gctggtctgt ggtgcatgtg tgtccctgtg atggcaatca
2161 ttgtagttgc tggccttcag aagaattgag gatctgatgg aggtttttta tgtatttatt
2221 ttctgttcac cttgtgaccc tgtgtcaaaa tttataaaga tacaaaaggc attactgaaa
2281 tggtactttc tgtaatttga tactatttgg cttaatcatc ttcacttgac tatttgtaat
2341 actgttgtaa tgttaactct gttaagtacc caagctgctt gtcttccacc aaagagtgct
2401 ttattaacaa gaatctgtga aaatcacatt taaacactgt tgcatgttgt aagaccaggt
2461 ggtaccttag taacctaaaa cttgcaagag aatattaatg gtagctttag aagactcagg
2521 aggagaaact gacttcagag ttggaagatg ttgcaagtcg ttccttttc tgtccttcag
2581 ggactgaaga actgggaggc tgcccattgt ttggttgcca gtcatacaaa ttaaaatcat
2641 atttccttcc atgaatggaa gaaacacact attggttttt ccccttggaa acagcaatcc
2701 caaataatgt cggcttacaa aaaaaaaaag ttaccacttt tttagagtcc ttccctgtaa
2761 cattggattt tttttttccc ttatgagatc cacctaaggc cattgacgtg gcctgcgatc
2821 tcagtgacaa tgatctgctt ctggatctca ctgttgcctt tggttaggga acacagagtg
2881 cttctcccgc agccctactg gaacacagca gagtctgtgc catgaagcag ttacagaaac
2941 agaattgatg tgctgctaaa aaaaaaaaaa aaaatggggc ccgggggggc gtccgccggc
3001 cctgcgggcc gccggtgaaa taccactact ctgatcgttt tttcactgac ccggtgaggc
3061 gggggggcga gccccgaggg gctctcgctt ctggcgcg

FIGURE 9H. Homo sapiens CDC14 homolog B, Amino acid sequence, isoform 1 (SEQ ID NO: 16)

1 mkrkserrss waaappcsrr csstspgvkk irsstqqdpr rrdpqddvyl ditdrlcfai
61 lysrpksasn vhyfsidnel eyenfyadfg plnlamvyry cckinkklks itmlrkkivh
121 ftgsdqrkqa naaflvgcym viylgrtpee ayrilifget syipfrdaay gscnfyitll
181 dcfhavkkam qygflnfnsf nldeyehyek aengdlnwii pdrfiafcgp hsrarlesgy
241 hqhspetyiq yfknhnvtti irlnkrmyda krftdagfdh hdlffadgst ptdaivkefl
301 dicenaegai avhckaglgr tgtliacyim khyrmtaaet iawvricrpg svigpqqqfl
361 vmkqtnlwle gdyfrqklkg qengqhraaf sklls gvddi singvenqdq qepepysddd
421 eingvtqgdr lralksrrqs ktnaipltls isrtktvlr

FIGURE 9I. Homo sapiens CDC14 homolog B, Nucleic acid sequence, transcript variant 2 (SEQ ID NO: 17)

1 cacggaacag ccctcctggg gtccccacga gccgcgtcct gctgtgcccc ggcgcctacg
61 cagcagcggc cgcggccgcg gtgggcacgc acggttaccc cgggcagctc cggccgccag

121 ctgcagcccc gtcgcctcgg ccgcgccagc cggctgcggg cacctggggg cgggctgggg
181 gcgccggccg cggcaggagg cgctgtagcg agggctgcgg cgccggtcct gcggcggccg
241 cgggaggcag cggggcaggc gctgtgggcc gggctcctcc tccggctcct gcgcgaccgc
301 ctcccgccgg gctctgccgg cgcccgccgt ccccgcagcg ccgctctgcg cccgccgccc
361 cgagcgcccg cgcggggctg gcgggagcct cggcgggcgc gcgggcgcgc ggggccatgg
421 tcgtggcccc ctgacgggcc gcggccgcct ccatgaagcg gaaaagcgag cggcggtcga
481 gctgggccgc cgcgcccccc tgctcgcggc gctgctcgtc gacctcgccg ggtgtgaaga
541 agatccgcag ctccacgcag caagacccgc gccgccggga cccccaggac gacgtgtacc
601 tggacatcac cgatcgcctt tgttttgcca ttctctacag cagaccaaag agtgcatcaa
661 atgtacatta tttcagcata gataatgaac ttgaatatga gaacttctac gcagattttg
721 gaccactcaa tctggcaatg gtttacagat attgttgcaa gatcaataag aaattaaagt
781 ccattacaat gttaaggaag aaaattgttc attttactgg ctctgatcag agaaaacaag
841 caaatgctgc cttccttgtt ggatgctaca tggttatata tttgggggaga accccagaag
901 aagcatatag aatattaatc tttggagaga catcctatat tcctttcaga gatgctgcct
961 atggaagttg caatttctac attacacttc ttgactgttt tcatgcagta aagaaggcaa
1021 tgcagtatgg cttccttaat ttcaactcat ttaaccttga tgaatatgaa cactatgaaa
1081 aagcagaaaa tggagattta aattggataa taccagaccg atttattgcc ttctgtggac
1141 ctcattcaag agccagactt gaaagtggtt accaccaaca ttctcctgag acttatattc
1201 aatattttaa gaatcacaat gttactacca ttattcgtct gaataaaagg atgtatgatg
1261 ccaaacgctt tacggatgct ggcttcgatc accatgatct tttctttgcg gatggcagca
1321 cccctactga tgccattgtc aaagaattcc tagatatctg tgaaaatgct gagggtgcca
1381 ttgcagtaca ttgcaaagct ggccttggtc gcacgggcac tctgatagcc tgctacatca
1441 tgaagcatta caggatgaca gcagccgaga ccattgcgtg ggtcaggatc tgcagacctg
1501 gctcggtgat tgggcctcag cagcagtttt tggtgatgaa gcaaaccaac ctctggctgg
1561 aaggggacta ttttcgtcag aagttaaagg ggcaggagaa tggacaacac agagcagcct
1621 tctccaaact tctctctggc gttgatgaca tttccataaa tggggtcgag aatcaagatc
1681 agcaagaacc cgaaccgtac agtgatgatg acgaaatcaa tggagtgaca caaggtgata
1741 gacttcgggc cttgaaaagc agaagacaat ccaaaacaaa cgctattcct ctcacagtaa
1801 ttcttcaatc cagtgttcag agctgtaaaa catctgaacc taacatttct ggcagtgcag
1861 gcattactaa aagaaccacc agatctgctt caaggaaaag cagtgttaaa agtctctcca
1921 tttcaaggac taaaacagtc ttgcgttaag taaaaacctg tgaccagagc tgaaggaaga
1981 ctctaggact gaaaactgca acagaaatta gcacaatttg aaaacaaaac aaaattgcaa
2041 aagccttagt tgctttttcc acctaagaag ttgatcaatg gagaaaatgt ccactggagt
2101 ttgaataatg aactttgagt ttgggtgcaa gcaaatgact cagagaaggg tccagctctc
2161 aagctgaatg acaaacatgc tgttgtaaat ttagtctcag gtgtaaatac ccaagccctc
2221 tggtacccag ggagctggct ggtctgtggt gcatgtgtgt ccctgtgatg gcaatcattg
2281 tagttgctgg ccttcagaag aattgaggat ctgatggagg tttttatgt atttattttc
2341 tgttcacctt gtgaccctgt gtcaaaattt ataaagatac aaaaggcatt actgaaatgg
2401 tactttctgt aatttgatac tatttggctt aatcatcttc acttgactat ttgtaatact
2461 gttgtaatgt taactctgtt aagtacccaa gctgcttgtc ttccaccaaa gagtgcttta
2521 ttaacaagaa tctgtgaaaa tcacatttaa acactgttgc atgttgtaag accaggtggt
2581 accttagtaa cctaaaactt gcaagagaat attaatggta gctttagaag actcaggagg
2641 agaaactgac ttcagagttg gaagatgttg caagtcgttc ctttttctgt ccttcaggga
2701 ctgaagaact gggaggctgc ccattgtttg gttgccagtc atacaaatta aaatcatatt
2761 tccttccatg aatggaagaa acacactatt ggtttttccc cttggaaaca gcaatcccaa
2821 ataatgtcgg cttacaaaaa aaaaaagtta ccactttttt agagtccttc cctgtaacat
2881 tggatttttt ttttccctta tgagatccac ctaaggccat tgacgtggcc tgcgatctca
2941 gtgacaatga tctgcttctg gatctcactg ttgcctttgg ttagggaaca caactagtaa
3001 ctctgcagag tgccttctcc cgcagcccta ctggaacaca gcagagtctg tgccatgaag
3061 cagttacaga aacagaattg atgtgctgcc aaaaaaaaaa aaaaaatggg gcccgaaata
3121 aaagaatata tagtactcac ctcagttcct tccataagaa gtgggtggtt taatgattgt
3181 taagccattt ttgcctgtgc cgggagcatg gagggctgag atgtcgacag gcagtgggaa
3241 acaaatgccc tcctaagcca caaggcgtgc gccagattag taggcaactc cattttaaga
3301 agctgccttt ttcacaaaac tggaagaaat aaaagcggtt ggaataaaca agttaaaagt
3361 ctttaatgca aaaagtaatt gaaaggcagt gcctccattt tggtgtactt tcttggaaga
3421 aagtataaaa ttgaccggca tcatgagaga cggaagatgc cgtgttctca gccaaacaag
3481 caactctttc cccgccaggc actgtcgggt ggggtcaggc cagcttttaa acactgggga
3541 ctggatcaca gaaaaacagt ggttttctgt ccctggaaat gaataggcac aaagacccac

3601 ttggctgtgg gcagactact cttcaataag atttgggtgg gaggaggaac attccttttg
3661 ctattttgag ctgagacaat ataaatattc aaactgtgcc atgcataaag cattgaattc
3721 tcagggcacc tcttcttccc cttacccctt ttaaggccat cccctccatt aataataatc
3781 caggtagttg tgaaaatcgt gcttctatct gatcccttct tagtttggct tttcatccca
3841 tcagaacaag taaacgtagg cgccacagct cttgtgagta ctgtctccct cacggtgaat
3901 gagcctcctg gtgtttcgtc caagaaaaga aagggtgtca ctggaaccac agcccttttt
3961 cattttataa actgcctctt catgttgcct gctcaagttt ccacctagaa ttgctatcac
4021 tgtggctctt tctaaaaatc tttctattta actggttcac tgaaattagt catagaaaac
4081 ttgtgatttg gtgaagaggc attccttgta ataaccaaat gacttgggat ggtgtgcata
4141 gcaagggcag tgttacactt atgaggactg tctctagcat ccaggaagtc tctgggtctg
4201 agggatggaa agttcttcct gctatgaatg agagtggact cttcccctca cccccaactg
4261 aaaccacaaa caaccagaat cttctggaat tctgacttag agtcgttgtt atagaagacc
4321 ttgttgctat ggaacatgaa actgtgtgtc agatggagag atccccttaa cctaagagcc
4381 ttaaatagcc ctgaaagtac actgggacgg tttgcgatgg aattaaaatt ggaagtgaat
4441 attttaggt gctcttgaag ctttctgggg actcaaaatt atcaaaagtc agggacagtc
4501 cggaggaaga gcgtctgcaa aactgggttc ctagaagtat agacggactt agctttttgt
4561 agaatttggt gaggagcagc gcctcgtgag agcagaatgg cctggcgtgg ccagtgcttc
4621 ccgg

FIGURE 9J. **Homo sapiens CDC14 homolog B, Amino acid sequence, isoform 2 (SEQ ID NO: 18)**

1 mkrkserrss waaappcsrr csstspgvkk irsstqqdpr rrdpqddvyl ditdrlcfai
61 lysrpksasn vhyfsidnel eyenfyadfg plnlamvyry cckinkkkks itmlrkkivh
121 ftgsdqrkqa naaflvgcym viylgrtpee ayrilifget syipfrdaay gscnfyitll
181 dcfhavkkam qygflnfnsf nldeyehyek aengdlnwii pdrfiafcgp hsrarlesgy
241 hqhspetyiq yfknhnvtti irlnkrmyda krftdagfdh hdlffadgst ptdaivkefl
301 dicenaegai avhckaglgr tgtliacyim khyrmtaaet iawvricrpg svigpqqqfl
361 vmkqtnlwle gdyfrqklkg qengqhraaf skllsgvddi singvenqdq qepepysddd
421 eingvtqgdr lralksrrqs ktnaipltvi lqssvqsckt sepnisgsag itkrttrsas
481 rkssvkslsi srtktvlr

FIGURE 9K. **Homo sapiens CDC14 homolog B), Nucleic acid sequence, transcript variant 3 (SEQ ID NO: 19)**

1 cacggaacag ccctcctggg gtccccacga gccgcgtcct gctgtgcccc ggcgcctacg
61 cagcagcggc cgcggccgcg gtgggcacgc acggttaccc cgggcagctc cggccgccag
121 ctgcagcccc gtcgcctcgg ccgcgccagc cggctgcggg cacctggggg cgggctgggg
181 gcgccggccg cggcaggagg cgctgtagcg agggctgcgg cgccggtcct gcggcggccg
241 cgggaggcag cggggcaggc gctgtgggcc gggctcctcc tccggctcct gcgcgaccgc
301 ctcccgccgg gctctgccgg cgcccgccgt ccccgcagcg ccgctctgcg cccgccgccc
361 cgagcgcccg cgcggggctg gcgggagcct cggcgggcgc gcgggcgcgc ggggccatgg
421 tcgtggcccc ctgacgggcc gcggccgcct ccatgaagcg gaaaagcgag cggcggtcga
481 gctgggccgc cgcgcccccc tgctcgcggc gctgctcgtc gacctcgccg ggtgtgaaga
541 agatccgcag ctccacgcag caagacccgc gccgccggga cccccaggac gacgtgtacc
601 tggacatcac cgatcgcctt tgttttgcca ttctctacag cagaccaaag agtgcatcaa
661 atgtacatta tttcagcata gataatgaac ttgaatatga gaacttctac gcagattttg
721 gaccactcaa tctggcaatg gtttacagat attgttgcaa gatcaataag aaattaaagt
781 ccattacaat gttaaggaag aaaattgttc attttactgg ctctgatcag agaaaacaag
841 caaatgctgc cttccttgtt ggatgctaca tggttatata tttggggaga accccagaag
901 aagcatatag aatattaatc tttggagaga catcctatat tcctttcaga gatgctgcct
961 atggaagttg caatttctac attacacttc ttgactgttt tcatgcagta aagaaggcaa
1021 tgcagtatgg cttccttaat ttcaactcat ttaaccttga tgaatatgaa cactatgaaa
1081 aagcagaaaa tggagattta aattggataa taccagaccg atttattgcc ttctgtggac

1141 ctcattcaag agccagactt gaaagtggtt accaccaaca ttctcctgag acttatattc
1201 aatattttaa gaatcacaat gttactacca ttattcgtct gaataaaagg atgtatgatg
1261 ccaaacgctt tacggatgct ggcttcgatc accatgatct ttctttgcg gatggcagca
1321 cccctactga tgccattgtc aaagaattcc tagatatctg tgaaaatgct gagggtgcca
1381 ttgcagtaca ttgcaaagct ggccttggtc gcacgggcac tctgatagcc tgctacatca
1441 tgaagcatta caggatgaca gcagccgaga ccattgcgtg ggtcaggatc tgcagacctg
1501 gctcggtgat tgggcctcag cagcagtttt tggtgatgaa gcaaaccaac ctctggctgg
1561 aaggggacta ttttcgtcag aagttaaagg ggcaggagaa tggacaacac agagcagcct
1621 tctccaaact tctctctggc gttgatgaca tttccataaa tggggtcgag aatcaagatc
1681 agcaagaacc cgaaccgtac agtgatgatg acgaaatcaa tggagtgaca caaggtgata
1741 gacttcgggc cttgaaaagc agaagacaat ccaaaacaaa cgctattcct ctcaccgatg
1801 gttggctgtc ccaggctgtc acctttctag accggcttct gatctggctc gggatccaca
1861 aggactagac ctgcggggaa ggtctctcct ggacacgccc gttgcccact gcaagttctc
1921 tccaggtgca attgaagcct ctcagcagcg gaggccgcca tgtggagaga gcaggcaggc
1981 ccactgctgc tgagaacagg gcaggcacgg gcagctcctg ttctgccttt cccagcttcg
2041 gagacgcagg ctcagctgct ccgaagcacc tgccagcacc gcacagtaca gtttcagagg
2101 acagcagtct ccttcccgtg aagctcccat gtgctggaat ggcatggact tgctgatcaa
2161 cagaaggaaa tggtctgaag tctgaccagc acaaggaagg aggctggctg gctcagaggg
2221 gcccaccttg cgtggaatga aaacgccaaa ggctcatgag caacattagg ctagaggggt
2281 cttgttcaaa gcatccaact ctgacttcgg aggcattccc agccggcagc agtgtgtcca
2341 gcctgcctct tcccaggctg gtctgacatg cagcttaggc tttcatccca agttaggtac
2401 tgacccctcc ctcttgggca gcacctccct ttttaaaaaa atttttttt cttccaaaga
2461 cagagtcttg ctcttgttgt ccaggctgga gtgcagtggc gcgatctagg ctcactgcaa
2521 cctccttctc ccaggttcaa gcgactctcc tgcctcagcc tcctgagtag ctgggattat
2581 aggcgtctgc caccacgccc ggctaatttc tgtatttta gtagagacag ggtttcacca
2641 tgttggccag gctggtctcg aactcctgac ctcaagtgat ctgcctgcct tggcctccca
2701 aagtgctggg attacaggtg tcagccaccg cacccagcca agcaccccta tctctagagg
2761 atctggcccc ccagcccagt tactgcaggg cagctttccc cacctggtga caggctgtgc
2821 gcagcagccc caggacctca ccctgagctg agtcttcagg agccgccctg gtggcacaac
2881 tcagacaccc ctgaggccta gcagtcaact cctgattcag acatgatcca gtccagcctg
2941 ggcttggcta taaccagctc aaacttgctt gacctccact tttcaggaga cttggggacg
3001 acagccctca tcggcgtctt tcatggggtt aatctgcttg agtctaagtc gccagccaga
3061 aacgtggtgc ccagggtgcc ctgcctcagg acatgtccac acccacgtca caagcacctg
3121 aggagtccgg ccggggcact gtggtccaaa aggtcctgcc gcctccgcat ctgactgtcc
3181 caacggcatg ctggtgacac cccccctgccc ttcgcttctg tcctccctgg cttctctggg
3241 gcacttgggg ctatgtacaa cctggcacga tccagaaagg gtgcaaacaa aatgcctaca
3301 tccaggcaca cgaccaagtc agcgagagct agccctggta agcaaacata gcccattaca
3361 ggttcagaac gtgcaccggg ttccccaaaa ctgtcttcaa ccacatgact caacagctct
3421 atgggatagg aactgtcagt gtttttgcaa ctgcaacatt aaaccaagtg ctgtgggctt
3481 ttcaagtatt attcacaaca ctaaaggaaa gtttcttcaa agggctctct ggctaatctt
3541 caaagccgca gttaggcaaa atgacagtgt gacagcttca aagccactga ctcatgacac
3601 agccctgatg ttgtaccggc taggttcaga tttcagaaat cagggcactt gcatccattg
3661 ccttttccag gaaagggaag aaaacactca gttgataaac cttagtactc agataataaa
3721 taagagacca aaagtaggct atcacccaaa gcaaacatcc ttaactgacc ctaacgtgta
3781 tggattcaac tttgattatt caacaaaatc atgaccgact gctgtggcct ggagtaacca
3841 aaggactgtt ttctctacac aaagtcagga gcgaatacca acctttattt gcacttgggt
3901 tccagttcaa agccacctta gacagtgtgg caaagtggga aaaagcacag atcctgggac
3961 caaggttcag attccatctc aagcgagcat atgaactgtg tgacaacagg cagacagtac
4021 ctctgtgtct atgagaaagc ggggagagca acaccccagc ttctagcagc tctacagctg
4081 cctggacctg caggccctcc taggcccact tcctccccag cacagtgtgt gttcccgggc
4141 gtgtgtggct ctgggtccag ctctgttcag ggtgggactc caggtgaatt actgaacctc
4201 tgaggtgtac ccccaacccc aaactttcac caaaagcaat aaagaggaac tctagaactg
4261 gagccaggac taagtgagaa aaactgctta taagtgctta ataaatacta gttatttaca
4321 acttttgctc aagccgaggg cagaggcctt tgtacgcagc tgccgaactc tgactctagt
4381 tctgcggaag aaaaggatgc ggtatttgct tttgccatga tccctttcca tttgattggc
4441 aggttaaata acatggtttt tgaagtcaca tacttaatat tcttcctaaa aaccacccaa
4501 acactagatg tgtgtgtgca cacacacaga aaccacgagg tagtttaaat caccattaaa
4561 aatcaacgct ttctctgatt ctgtgtcaca gagtggtggc cagtggctac acaatagcga

4621 tttttaaatg attggttaag tgaaaaccag aactcaaaat attccaggag agaagataac
4681 atttacaagt aaacagtaag tgcaattgta ttttaatttc ttggtctccg aaaactcagc
4741 tgtgactgct ttccattaac agttccagct ctatgtgttt cctctaacgc taaaggcaca
4801 gcccccggga atctactgct tcctaagagt ctccatggag tctattttac aacctccttt
4861 ccctccatgc ttccgcggag gagtctatac tatctctata tacacatttt aaacattatt
4921 cttcatttga aattccttca ataaaaacac agtcaccatt aaaaaaaa

FIGURE 9L. Homo sapiens CDC14 homolog B), Amino acid sequence, isoform 3 (SEQ ID NO: 20)

1 mkrkserrss waaappcsrr csstspgvkk irsstqqdpr rrdpqddvyl ditdrlcfai
61 lysrpksasn vhyfsidnel eyenfyadfg plnlamvyry cckinkklks itmlrkkivh
121 ftgsdqrkqa naaflvgcym viylgrtpee ayrilifget syipfrdaay gscnfyitll
181 dcfhavkkam qygflnfnsf nldeyehyek aengdlnwii pdrfiafcgp hsrarlesgy
241 hqhspetyiq yfknhnvtti irlnkrmyda krftdagfdh hdlffadgst ptdaivkefl
301 dicenaegai avhckaglgr tgtliacyim khyrmtaaet iawvricrpg svigpqqqfl
361 vmkqtnlwle gdyfrqklkg qengqhraaf skllsgvddi singvenqdq qepepysddd
421 eingvtqgdr lralksrrqs ktnaipltdg wlsqavtfld rlliwlgihk d

FIGURE 9M. Homo sapiens hypothetical protein MGC26484 (MGC26484), Nucleic acid sequence (SEQ ID NO: 21)

1 agcggggcgg cctccaggaa gcggaaaagc aaggggcggt cgagctgggc cgccgcgccc
61 cactgctcgc cgcgctgctc tttgacctcg cagggtgtga agaagatgcg cagctccacg
121 ctgcaggacc cgcgccgccg ggacccccag gacgacgtgt acctggacat caccgatcgc
181 cttcgttttg ccattctcta cagcagacca aagagtgcat caaatgtaca ttatttcagc
241 atagataatg aactcgaata tgagaacttc tccgaagact ttggaccact caatctggca
301 atggtttaca gatattgttg caagataaat aagaaattaa agtccattac aatgttaagg
361 aagaaaattg ttcattttac tggctctgat cagagaaaac aagcaaatgc tgccttcctt
421 gttggatgct acatggttat atacttgggg agaaccccag aagcagcata tagaatatta
481 atctttggag atacacccta tattcctttc agagatgctg cctatggaag ctgcaatttc
541 tacattacac ttcttgactg ttttcatgca gtaaagaagg caatgcagta tggcttcctt
601 aatttcaact catttaacct tgatgaatat gaacactatg aaaaagcaga aaacggagat
661 ttaaattgga taataccaga ccgatttatt gccttctgtg gacctcattc aagagccaga
721 cttgaaagtg gttaccacca acattctccc gagacttata ttcaatattt taagaatcac
781 aatgttacta ccattattcg tctgaataaa aggatgtatg atgccaaacg ctttacggat
841 gctggcttcg atcaccatga tcttttcttt gcggatggca gcacccctac tgatgccatt
901 gtcaaaagat ttctggatat ctgtgaaaat gctgagggtg ccattgcagt acattgtaaa
961 gctggccttg gtcgcacagg cactctgata gcctgctaca tcatgaagca ttacaggatg
1021 acagcagccg agaccattgc gtgggtaagg atctgcagac ctggcttggt gatcgggcct
1081 cagcagcagt ttttggtgat gaagcaaaca agcctctggc tggaagggga ctatttcgt
1141 cagaggttaa aggggcagga gaatggacaa cacagagcag ccttctccaa acttctctct
1201 ggtgttgatg acatttccat aaatggggtc gagaatcaag accagcaaga acccaaacct
1261 tacagtgatg acgacgaaat caatggagtg acacaaggtg atagaagtcg ggccctgaaa
1321 aggcgaagac aatcaaaaac aaacgatatt cttctcccat ctcccctggc tgtgctgacc
1381 tttacactgt gtagtgttgt catctggtgg attgtttgtg actacattct tcccatcctg
1441 ctattctgac tcaaagattt cagaacatag tagccatcag gaccccctga cagatagctg
1501 cttctctcca tttcaaggac taaaatggtt ttgcattaag taaaaacctg tgaccagaac
1561 tgaaggaaga ctctaggaac tgaaaactgc aacagaaatt agcacaattt gaaaacaaaa
1621 caaaattgca aaagacttag ttgctttcca cctaagaagc taatcaatgg ggaaaatgtc
1681 cacggggttt cagtaatgaa cttttgagtt tgggtgcaag caaatgtacc aagaccagct
1741 cagtcaggga gaccctaccc cagtggtgct agaggaatta aagacacaca cacagaaata
1801 cagaggtgta aagtgggaaa tcaggggtct cacagccttc agagctgaga gccccgaaca
1861 gagatttacc cacatattga ttaacagcaa acaagtcatt agcattgttt ctatagatat

1921 taaattaact aaaagtatcc cttatgggaa acgaagggat gggcctaatt aaaggcatag
1981 gttgggctag ttaactgcgg caggagcacg tccttaaggc acagatggct catgctattg
2041 tttgtggctt aagaatgcct ttaagcagtt ttccactctg ggctgggtgg gccaggtgtt
2101 ccttgccctc attccggtaa acccacaccc ttccagcgtg ggcattaggg ccattatgaa
2161 catgttacag tgctgcagag attttattta tggccagttt tggggccagt ttatggctgg
2221 atttgggggg gcttgctccc aaaaagaatg actcaaagaa aggcccagct ctcaagctga
2281 atgacaaaaa tgctgttgta aatttagtct caggtgtaaa tacccaagcc ctctggtact
2341 cagggagctg gctggtctgt ggtgcatgtg tgtccttgtg atggcaagca ttgtagttga
2401 tggccttcag aagaattgag gatctgatgg aggtttttta atgtatttat tttctgttca
2461 gcttgtgacc ctgtgtcaaa atttgtaaag atacaaaagg cattactgaa atggtacttt
2521 ctgtaatttg atactatttg gctttaccat cttcacttga ctgtttgtaa tactgtagta
2581 atattaactc tgataagtac ccaagctgct tgtcttctac caaagagtgc tttattaaca
2641 agaatctgtg aaaaccacat tttaaacgct gttgcatgtt gtaataccag gtgatacctt
2701 ggtaacctaa aacttgcaag agaatattag tggtagcttt agaagactca ggaggagaaa
2761 ctgacttcag agctggaaga tgttgcaagt cattcctttt tctgtccttt agggactgaa
2821 gaactgggag gttgcccatt gtttggttgc cagtcataca aattaaaatc atatttcctt
2881 ccatgcaaaa aaaaaaaaaa a

FIGURE 9N. Homo sapiens hypothetical protein MGC26484 (MGC26484), Amino acid sequence (SEQ ID NO: 22)

1 mrsstlqdpr rrdpqddvyl ditdrlrfai lysrpksasn vhyfsidnel eyenfsedfg
61 plnlamvyry cckinkklks itmlrkkivh ftgsdqrkqa naaflvgcym viylgrtpea
121 ayrilifgdt pyipfrdaay gscnfyitll dcfhavkkam qygflnfnsf nldeyehyek
181 aengdlnwii pdrfiafcgp hsrarlesgy hqhspetyiq yfknhnvtti irlnkrmyda
241 krftdagfdh hdlffadgst ptdaivkrfl dicenaegai avhckaglgr tgtliacyim
301 khyrmtaaet iawvricrpg lvigpqqqfl vmkqtslwle gdyfrqrlkg qengqhraaf
361 skllsgvddi singvenqdq qepkpysddd eingvtqgdr sralkrrrqs ktndillpsp
421 lavltftlcs vviwwivcdy ilpillf

FIGURE 9O. Homo sapiens similar to CDC14 homolog B, isoform 3, Nucleic acid sequence (SEQ ID NO: 23)

1 ggccgcctcc aggaagcgga aaagcaagcg gcggtcgagc tgggccgccg cgcccctctg
61 ctggcagcgc tgctctttga ccttgcaggg tgtgaagaag atgcgcagct ccacgctgca
121 ggacccgcgc cgctgggacc cccaggacga cgtgtacctg gacatcaccg atcgccttcg
181 ttttgccatt ctctacagca gaccaaagag tgcatcaaat gtacattatt tcagcataga
241 taatgaactc gaatatgaga acttctccga agactttgga ccactcaatc tggcaaatgg
301 tttacagata ttgttgcaag ataaataaga aattaaagtc cattacaatg ttaaggaaga
361 aaattgttca ttttactggc tctgatcaga gaaaacaagc aaatgctgcc ttccttgttg
421 gatgctacat ggttatatac ttggggagaa ccccagaagc agcatataga atattaatct
481 ttggagatac atcctatatt cctttcagag atgctgccta tggaagctgc aatttgtaca
541 ttacacttct ttcctgtttt catgcagtaa agaaggcaat gcagtatggc ttccttaatt
601 tcaactcatt taaccttgat gaatatgaac actatgaaaa agcagaaaac ggagatttaa
661 attggctaat accagaccga tttattgcct tctgtggacc tcattcaaga gccagacttg
721 aaagtggtta ccaccaacat tctcccgaga cttatattca atattttaag aatcacaatg
781 ttactaccat tattcgtctg aataaaagga tacatgatgc caaacgcttt acggatgctg
841 gcttcgatca ccatgatctt ttctttgcag atggcagcac ccctactgat gccattgtca
901 aaagatttct ggatatctgt gaaaatgctg agggtgccat tgcagtacat tgtaaagctg
961 gccttggtcg cacaggcact ctgatagcct gctacatcat gaagcattac aggatgacag
1021 cagccgagac cattgcgtgg gtaaggatct gcagacctgg cttggtgatc gggcctccgc
1081 agcagttttt ggtgatgaag caaacaagcc tctggctgga aggggactat tttcgtcaga
1141 agttaaaggg gcaggagaat ggacaacaca gagcagcctt ctccaaactt ctctctggtg
1201 ttgatgacat ttccataaat ggggtcgaga atcaagacca gcaagaaccc aaaccttacg

```
1261 gtgatgacga cgaaatcaat ggagtgacac aaggtgatag aagtcgggcc ctgaaaaggc
1321 aaagacaatc aaaaacaaac gatattcttc tcccatctcc cctggctgtg ctgaccttta
1381 cactgtgtag tgttgtcatc tggtggattg tttgtgacta cattcttccc atcctgctat
1441 cctgactcga agatttcaga acatagtagc catcaggacc ccctgacaga tagctgcttc
1501 tctccatttc aaggactaaa actgttttgc attaagtaaa aacctgtgac cagaactgaa
1561 ggaagactct aggaactgaa aactgcaaca gaaattagca caatttgaaa acaaaacaaa
1621 attgcaaaag acttagttgc tttccaccta agaagctaat caatggagaa aatgtccact
1681 ggggtttcag taatgaactt ttgagtttgg gtgcaagcaa atgtaccaag accagctcag
1741 tcagggagac cctaacccag tggtgctaga ggaattaaag acacacacac agaaatacag
1801 aggtgtaaag tgggaaatca ggggtctcac agccttcaga gctgagagac ccgaacagag
1861 atttacccac atattgatta acagcaaacc agtcattagc attgtttcta tagatattaa
1921 ataaactaaa agtatccctt atgggaaacg aagggatggg cctaattaaa ggcataggtt
1981 gggctagtta actgcggcag gagcacgtcc ttaaggcaca gatcgctcat gctattgttt
2041 gtggcttaag aatgccttta agcagttttc cactctgggc tgggtgggcc aggtgttcct
2101 tgccctcatt ccggtaaacc cacacccttc cagcgtgggc attagggcca ttatgaacat
2161 gttacagtgc tgcagagatt ttgtttatgg ccagttttgg ggccagttta tggctggatt
2221 ttgggggctt gctcccaaaa agaatgactc aaagaaaggc ccagctctca agctgaatga
2281 caaaaatgct gttgtaaatt tagtctcagg tgtaaatacc caagccctct ggtacccagg
2341 gagctggctg gtctgtggtg catgtgtgtc cttgtgatgg caagcattgt agttgctggc
2401 cttcagaaga attgaggatc tgatggaggt tttttatgta tttattttct gttcagcttg
2461 cgaccctgtg tcaaaatttg taaagataca aaaggcatta ctgaaatggt actttctgta
2521 atttgatact atttggcttt atcatcttca cttgactgtt tgtaatactg tagtaatatt
2581 aactctgata agtacccaag ctgcttgtct tccaccaaag agtgctttat taacaagaat
2641 ctgtgaaaac cacatttaaa cactgttgca tgttgtaata tcaggtgata ccttggtaac
2701 ctaaaacttg caagagaata ttaatggtag ctttagaaga ctcaggagga gaaactgact
2761 tcagagttgg aagatgttgc aagtcattcc tttttctgtc ctttagggac tgaagaactg
2821 ggaggttgcc cattgtttgg ttgccagtca tacaaattaa aatcatattt ccttccatgc
```

FIGURE 9P. Homo sapiens similar to CDC14 homolog B, isoform 3, Amino acid sequence (SEQ ID NO: 24)

```
  1 mnsnmrtspk tldhsiwqmv yrycckinkk lksitmlrkk ivhftgsdqr kqanaaflvg
 61 cymviylgrt peaayrilif gdtsyipfrd aaygscnlyi tllscfhavk kamqygflnf
121 nsfnldeyeh yekaengdln wlipdrfiaf cgphsrarle sgyhqhspet yiqyfknhnv
181 ttiirlnkri hdakrftdag fdhhdlffad gstptdaivk rfldicenae gaiavhckag
241 lgrtgtliac yimkhyrmta aetiawvric rpglvigppq qflvmkqtsl wlegdyfrqk
301 lkgqengqhr aafskllsgv ddisingven qdqqepkpyg dddeingvtq gdrsralkrq
361 rqsktndill psplavltft lcsvviwwiv cdyilpills
```

FIG. 10B: Real-time PCR mediated analysis of NR1D1 expression in 3T3-L1 cells differentiated from preadipocytes to mature adipocytes (DCt (d0) = 36)

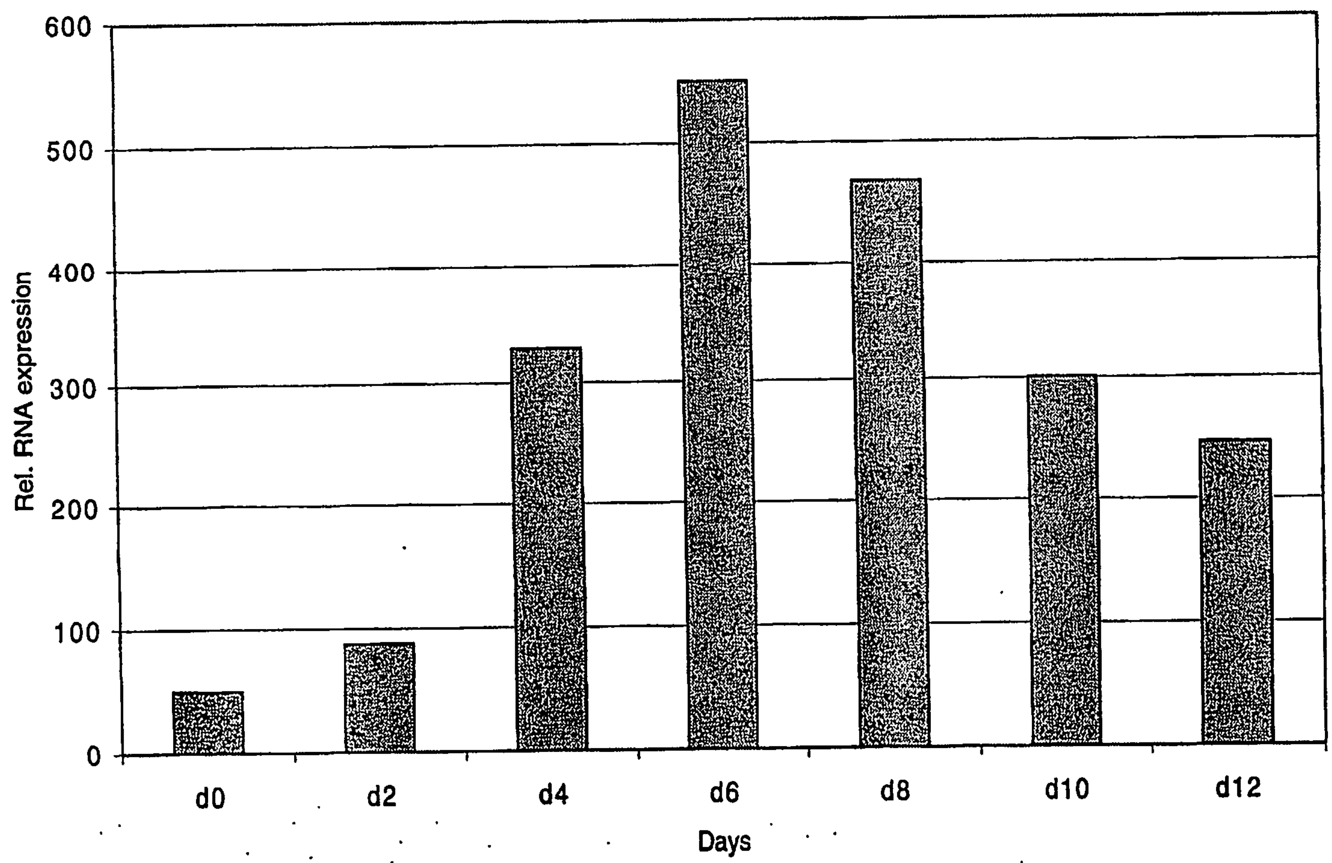
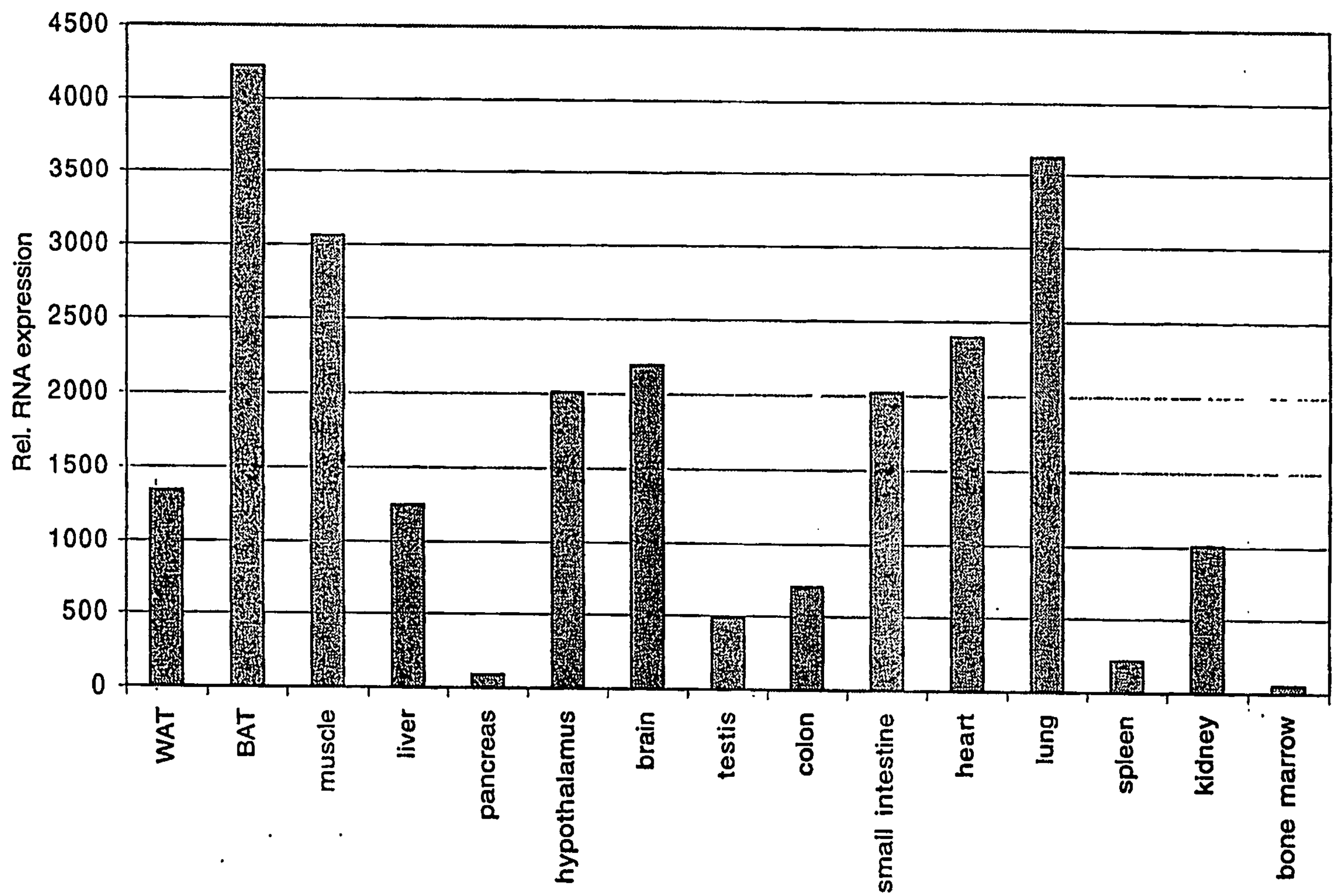


FIGURE 10: Expression of Tyrosine-protein kinase NR1D1 in mammalian tissues

FIG. 10A: Real-time PCR analysis of NR1D1 expression in wildtype mouse tissues (DCt (pancreas) = 36)



THIS PAGE BLANK (USPTO)

PCT Application
PCT/EP2003/014344